Engineered Alcohol Dehydrogenases for Stereoselective Chemical Transformations

THILAK REDDY ENUGALA
Enzymes are biomolecules built from amino acids and catalyze the chemical transformations in a cell. Enzymes are by nature stereoselective, biodegradable, environmentally friendly, and can perform catalysis in aqueous solutions and at ambient temperatures. Due to these advantages the use of enzymes as biocatalysts for chemical transformations has emerged as an attractive “greener” alternative to conventional chemical synthesis strategies. And, if naturally occurring enzymes cannot carry out the desired chemical transformations, the functional properties of enzymes can be modified by directed evolution or protein engineering techniques. Since enzymes are genetically encoded they can be optimized for desired traits such as substrate selectivity or improved catalytic efficiency. Considering these advantages and also keeping the synthetic and industrial application in mind, we have employed alcohol dehydrogenase-A (ADH-A) from Rhodococcus ruber DSM 44541 as a study object in engineering for new catalytic properties. ADH-A tolerates water miscible organic solvents, accepts a relatively wide range of aromatic sec-alcohols/ketones as substrates and is therefore a potentially useful biocatalyst for asymmetric synthesis of organic compounds.

Presented research work in this thesis has been primarily focused on engineering of ADH-A and characterization of resulting enzyme variants. The engineering efforts have aimed for altered substrate scope, as well as stereo- and regioselectivities. Furthermore, possible substrate promiscuity in engineered enzyme variants has also been addressed. In short, i).

Paper I: three sub sites, each consisting of two-three amino acid residues within the active-site cavity were exposed to saturation mutagenesis in step-wise manner, coupled to an in vitro selection for improved catalytic activity with the unfavored (R)-1-phenylethanol. The observed stereoselectivity could be explained partly by a shift in nonproductive substrate binding. ii).

Paper II is aimed specifically towards the improving the catalytic activity with aryl-substituted vicinal diols, such as (R)-1-phenylethane-1,2-diol, and the possibility to link the ADH-A reaction with a preceding epoxide hydrolysis to produce the acyloin 2-hydroxyacetophenone from rac-styrene oxide. iii).

Paper III is mainly focused towards studies of regioselectivity. Here, ADH-A and engineered variants were challenged with a substrate containing two sec-alcohol functions and the cognate di-ketone. The regioselectivity in wild type as well as in engineered variants could in part be explained by a combination of experimental and computer simulations. iv).

Paper IV is focused on elucidating possible effects on substrate promiscuities in engineered variants as compared to the wild type parent enzyme, when challenged with a spectrum of potential previously untested substrates.

Keywords: alcohol dehydrogenase-A, biocatalysts, protein engineering, enzyme kinetics, sec-alcohols, ketones, stereoselectivity, regioselectivity, substrate selectivity and promiscuity.
I dedicate this to

my Mom, Dad and rest of Family,

my Guru,

&

Bharat 🇮🇳 Sverige 🇸🇪
List of Papers

This thesis is based on the following research articles, which are referred to in the text by their Roman numerals.


† D.M. and T.R.E. contributed equally to this work.

Note: Reprints were made with permission from the respective publishers.
The following research articles were excluded from this Doctoral thesis.


Contribution Info

My contribution to the following research articles.

I. Purified the isolated library hits. Performed steady state and pre-steady state enzyme kinetics. Contributed to the writing and manuscript preparation.

II. Purified the isolated library hits. Performed steady state and pre-steady state enzyme kinetics. Contributed to the writing and manuscript preparation.

III. Purified the isolated library hits. Performed steady state and pre-steady state enzyme kinetics. Set-up the enzymatic reactions, purified and characterized reaction products by NMR and chiral reverse phase HPLC. Contributed to the writing and manuscript preparation.

IV. Purified the enzymes. Performed the steady state enzyme kinetics. Set-up the enzymatic reactions, purified and characterized reaction products by chiral reverse phase HPLC. Contributed to the writing and manuscript preparation.
Contents

Introduction ......................................................................................................................13
  Sustainable (Green) Chemistry .................................................................14
Enzymes ...........................................................................................................................15
  Enzyme Classification ..................................................................................16
  Chirality .............................................................................................................16
  Enzymes’ Role in Catalysis ........................................................................18
Biocatalysis ...................................................................................................................19
  Kinetic Resolution ..........................................................................................20
Industrial Biocatalysis and its Applications ...................................................21
  Traditional Limitations of Enzymes as Biocatalysts ......................................22
Enzyme Engineering ...............................................................................................22
  Directed Evolution .........................................................................................23
  Rational Design .............................................................................................24
  Semi-Rational Design/Iterative Saturation Mutagenesis ...............................24
Screening and Selection .............................................................................................26
Characterization .........................................................................................................27
  Enzyme Kinetics .............................................................................................27
  Enzyme Substrate Specificity and Selectivity ..............................................32
  Structure-Function Relationship ..................................................................34
Alcohol Dehydrogenases .........................................................................................35
  Structural Similarities of ADHs and Reaction ..............................................36
  Alcohol Dehydrogenase from Rhodococcus ruber .......................................37
ADH-A – Present investigations .............................................................................40
  Paper I .............................................................................................................40
  Paper II ..........................................................................................................50
  Paper III .........................................................................................................54
  Paper IV .........................................................................................................61
Concluding Remarks and Future Perspective ......................................................65
Svensk Sammanfattning ..........................................................................................67
Acknowledgements .................................................................................................69
References .................................................................................................................71
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH-A</td>
<td>Alcohol dehydrogenase A – <em>Rhodococcus ruber</em> DSM 44541</td>
</tr>
<tr>
<td>ADH</td>
<td>Alcohol dehydrogenase</td>
</tr>
<tr>
<td>B-fit</td>
<td>B-factor iterative test</td>
</tr>
<tr>
<td>CAST</td>
<td>Combinatorial active site test</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Cryo-electron microscopy</td>
</tr>
<tr>
<td>DE</td>
<td>Directed evolution</td>
</tr>
<tr>
<td>E</td>
<td>Enzymes</td>
</tr>
<tr>
<td>ee</td>
<td>Enantiomeric excess</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>epPCR</td>
<td>Error-prone polymerase chain reaction</td>
</tr>
<tr>
<td>Fe</td>
<td>Iron</td>
</tr>
<tr>
<td>FucO</td>
<td><em>E. coli</em> Propanediol oxidoreductase</td>
</tr>
<tr>
<td>FSA</td>
<td>D-Fructose 6-phosphate aldolase</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-performance liquid chromatography</td>
</tr>
<tr>
<td>ISM</td>
<td>Iterative saturation mutagenesis</td>
</tr>
<tr>
<td>KIE</td>
<td>Kinetic isotope effect</td>
</tr>
<tr>
<td>KR</td>
<td>Kinetic resolution</td>
</tr>
<tr>
<td>NAD(P)H</td>
<td>Nicotinamide adenine dinucleotide (phosphate)</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>R. ruber</td>
<td><em>Rhodococcus ruber</em></td>
</tr>
<tr>
<td>S</td>
<td>Substrate</td>
</tr>
<tr>
<td>SM</td>
<td>Saturation mutagenesis</td>
</tr>
<tr>
<td>StEH1</td>
<td><em>Solanum tuberosum</em> epoxide hydrolase 1</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>US(A)</td>
<td>United States (America)</td>
</tr>
<tr>
<td>Zn</td>
<td>Zinc</td>
</tr>
</tbody>
</table>
Twenty Amino Acids

**Set A: Amino Acids with Electrically Charged Side Chains**

- **Positive**
  - Arginine
  - Histidine
  - Lysine

- **Negative**
  - Aspartic acid
  - Glutamic acid

**Set B: Amino Acids with Polar Uncharged Side Chains**

- Serine
- Threonine
- Asparagine
- Glutamine

**Set C: Special Amino Acids**

- Glycine
- Cysteine
- Proline

**Set D: Amino Acids with Hydrophobic Side Chains**

- Alanine
- Valine
- Leucine
- Isoleucine
- Methionine
- Phenylalanine
- Tyrosine
- Tryptophan
Introduction

According to the 2019 revision of United Nations population prospect report, the current estimated World population is 7.7 billion. The growing population is expected to increase by an additional two billion people in the next 30 years, and could peak at nearly 11 billions by 2100.\(^1\) Apparently, in 1955 the UN estimated world population was 2.6 billion.\(^1\) This phenomenal growth to date of 5.1 billions has been driven largely by the increasing number of people that survive beyond the puberty age, and also accompanied by changes in fertility rates, increasing urbanization and migration. On the other hand, climate changes where the ice caps are melting, will lead to substantial rise of ocean levels. Because of human industrial activities the increased atmospheric concentration of CO\(_2\) and CH\(_4\) will increase, leading to a greenhouse effect.\(^2\) According to UN lead Intergovernmental Panel on Climate Change the current emission rates could increase the temperature up to 1.5 °C by 2030. This has been designated as a dangerous level, which can potentially cause harmful effects on biodiversity and livelihood.\(^2\)

The above mentioned trends have far-reaching implications for generations to come.\(^1,2\) One of the very important reasons for the overall population growth is due to an increased availability and accessibility of medicines, an improved quality of life and healthier food habits. The prolonged life expectancy, however, does not mean that people nowadays get less sick than in the past, and effective drugs are not accessible for everyone. Therefore, development of new pharmaceuticals as well as new chemicals will remain essential for our life commodities. Chemicals for hygiene and cosmetics, food additives, fertilizers, pesticides, herbicides, plastics, nylon, rubber, paints, refrigerants, organometallics, polymer, liquid crystal display in smartphones and TVs will continue to be expected and need to be developed.

The current way of life also comes with a high price in the form of hazardous chemical waste.\(^3,4\) The ever-increasing demand for chemicals, may also lead to an increased dependence on petroleum based products as feedstock for synthesis of new demanded chemicals. Many of which are produced by harsh conditions such as high temperatures, extreme pH, requiring hazardous heavy metal reagents and toxic organic solvents. Because of this, many of the manufactured chemicals and life supportive commodity products are endangering our planets’ environment and ecosystems. Decomposition of waste is often
expensive and can be time consuming.\textsuperscript{3,4} Thus, to both maintain our quality of life and at the same time make it possible for developing countries to increase theirs, and to do so in accordance with the UN Sustainable Development Goals (Table 1),\textsuperscript{1} societies must prevent buildup of hazardous waste. This can be addressed by applying a 6 \textit{Rs}-approach, which is short for: \textit{Rethink}, \textit{Reduce}, \textit{Reuse}, \textit{Recycle}, \textit{Refuse} and \textit{Repair}.\textsuperscript{5}

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
UN Goals & Descriptive labels \\
\hline
1 & No Poverty \\
2 & Zero Hunger \\
3 & Good Health \\
4 & Quality Education \\
5 & Gender Equality \\
6 & Clean Water and Sanitation \\
7 & Affordable and Clean Energy \\
8 & Decent Work and Economic Growth \\
9 & Industry, Innovation and Infrastructure \\
10 & Reduced Inequalities \\
11 & Sustainable Cities and Communities \\
12 & Responsible Consumption and Production \\
13 & Climate Action \\
14 & Life Below Water \\
15 & Life on Land \\
16 & Peace, Justice and Strong Institutions \\
17 & Partnerships for the Goals \\
\hline
\end{tabular}
\caption{UN Sustainable Development Goals}
\end{table}

Considering, the above mentioned goals,\textsuperscript{1} the chemicals industry must move away from petroleum-based approaches and improve existing and develop new sustainable production methods. Unfortunately, there are no straight-forward solutions available. Therefore, this issue must be tackled from several directions and at different levels. Hence, chemical industries are looking particularly for sustainable or greener alternatives, and cheaper solutions for their applications.\textsuperscript{6}

**Sustainable (Green) Chemistry**

The term Green Chemistry was coined by Paul Anastas in 1995. He formulated how chemical processes which carry environmental burden can be replaced with less polluting or non-polluting alternatives.\textsuperscript{6} Subsequently, Anastas and John Warner introduced the Twelve Principles of Green Chemistry\textsuperscript{7} (Table 2) as a tool to design new chemical processes and renewable production by considering all aspects of the synthesis process, such as environmental and handling toxicity, water pollution, feedstock and process cost.\textsuperscript{8,9} The Twelve
Principles of Green Chemistry emphasizes the UN Sustainable Development Goal 9.

**Table 2. The Principles of Green Chemistry**

<table>
<thead>
<tr>
<th>Green Principles</th>
<th>Descriptive labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Prevention of chemical waste</td>
</tr>
<tr>
<td>2</td>
<td>Atom efficiency/economy</td>
</tr>
<tr>
<td>3</td>
<td>Less toxic/hazardous chemical synthesis</td>
</tr>
<tr>
<td>4</td>
<td>Design and synthesize safer chemicals</td>
</tr>
<tr>
<td>5</td>
<td>Utilize safer solvents for synthesis</td>
</tr>
<tr>
<td>6</td>
<td>Energy efficient process design</td>
</tr>
<tr>
<td>7</td>
<td>Utilize renewable feedstocks</td>
</tr>
<tr>
<td>8</td>
<td>Avoid derivatization</td>
</tr>
<tr>
<td>9</td>
<td>Catalysis</td>
</tr>
<tr>
<td>10</td>
<td>Design the products for easy decomposition</td>
</tr>
<tr>
<td>11</td>
<td>Analytical methods for pollution prevention</td>
</tr>
<tr>
<td>12</td>
<td>Inherently safer chemical processes</td>
</tr>
</tbody>
</table>

These twelve illustrative principles strive for sustainable chemical synthesis at the molecular level. Apart from pharmaceutical industry, these principles have been applied to many industry sectors. For example, aerospace, automobile, cosmetic, electronics, energy, household products and economically competitive technologies. Recently, they were further summarized into the acronyms, PRODUCTIVELY and IMPROVEMENTS. To achieve a fully sustainable chemical process, preferably one should implement and execute all the principles. Some of the implementations include the use of aqueous solvents, ionic liquids, supercritical fluids, micro fluidics, and apply catalytic processes, including enzyme catalysis, rather than stoichiometric amounts.

**Enzymes**

Enzymes are biomolecules that catalyze the chemical reactions in cells. The existence of enzymes that can perform such diverse tasks is a testament to the power of natural evolution. Although, even before specific knowledge of enzymes were known, their catalytic abilities have been utilized by humans since ancient history. In food processing for example, curd (dahi or yoghurt), butter (ghee), cheese (paneer), and fermentation of carbohydrates for items such as idli, dosa, bread and beverages around the world. For example, in ~ 6000 BC Mesopotamia people fermented wine from sweet fruits. At the present date we know that enzymes in *Saccharomyces cerevisiae*, “baker’s yeast” carries out anaerobic oxidation of carbohydrates into ethanol. The French chemist Anselme Payen first reported the enzymes diastases in 1833, and a few years later Louis Pasteur postulated that fermentation of sugar into alcohol caused
by yeast cells, which he described as *ferments*.\textsuperscript{19,20} The term, ENZYME was coined by German physiologist Wilhelm Kühne\textsuperscript{21} in 1877, and two decades later fellow German chemist Eduard Buchner discovered that nonliving yeast can ferment sugar to ethanol, and he named this property as *zymase*.\textsuperscript{16}

In 1930’s Bayer AG used *S. cerevisiae* to form a precursor to ephedrine, L-phenylacetylcarninol also called as *R*-PAC by whole-cell biotransformation of benzaldehyde and D-glucose.\textsuperscript{22} Thereafter, the use of enzymes for chemical transformations slowly but steadily gained momentum. For example, the Danish company Chr. Hansen has an impressive foot-hold on dairy products related proteases, such are rennet (chymosin), lactases, catalases and lipases. And so, ENZYMOMOLOGY field opened-up the Pandora’s box for new possibilities and opportunities that can address various problems include to chemicals production that are endangering the planet and its ecosystems.

### Enzyme Classification

Enzymes are classified into six distinctive major classes (*Table 3*) by the type of the reaction they catalyze.\textsuperscript{16}

*Table 3. Enzyme classification (EC).*

<table>
<thead>
<tr>
<th>Class No</th>
<th>Class name</th>
<th>Catalyzed reaction types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oxidoreductases (EC 1)</td>
<td>Oxidation and Reduction reactions - electron transfer reaction C-H, C-C, C=X bonds</td>
</tr>
<tr>
<td>2</td>
<td>Transferases (EC 2)</td>
<td>Group transfer reactions - amino, acyl, sugar, phosphoryl, methyl, etc.</td>
</tr>
<tr>
<td>3</td>
<td>Hydrolases (EC 3)</td>
<td>Hydrolysis reactions - cleavage of acyl, alkyl, nitrile, C-R bonds using H₂O</td>
</tr>
<tr>
<td>4</td>
<td>Lyases (EC 4)</td>
<td>Addition or elimination reactions - on C=C, C=N and C=O bonds</td>
</tr>
<tr>
<td>5</td>
<td>Isomerases (EC 5)</td>
<td>Isomerization reactions - racemization, epimerization, rearrangement</td>
</tr>
<tr>
<td>6</td>
<td>Ligases (EC 6)</td>
<td>Formation or cleavage of bonds - of C-C, C-N, C-O, C-S bonds</td>
</tr>
</tbody>
</table>

### Chirality

In nature, molecules may exist in structural configurations that are either symmetric or asymmetric. Symmetric molecules cannot be distinguished from their mirror image objects while asymmetric molecules *cannot be superposed* onto their mirror images (*Figure 1*), like our left and right hands. This property is referred to as chirality.\textsuperscript{23} Chirality is derived from the Greek word for hand, and the two mirror images are referred to as *enantiomers*. A mixture of two enantiomers in a 1:1 ratio is a *racemate*. The enantiomers in a racemic mixture behave identically in an achiral environment, but when subjected to a likewise
asymmetric (chiral) environment or exposed to plane-polarized light, enantiomers act, and rotate light, differently. The existence of enantiomers has been known to humans since the mid 19th century; in 1848 Louis Pasteur treated an aqueous solution of racemic tartrate salt with *Penicillium glaucum*, which resulted in the enrichment of (-)-tartaric acid due to the consumption of (+)-tartaric acid by the fungus; a kind of whole-cell enzyme kinetic resolution. The term kinetic resolution will be described below.

**Figure 1.** The non-superimposable mirror images of *Eldorado* – and *Ambassadeurs*, are lithography prints depicted by artist Henri Toulouse-Lautrec in 1870.

Most biomolecules are built from chiral building blocks, such as proteins and polysaccharides which are made from L-amino acids and D-carbohydrates, respectively. Therefore, enzymes which are proteins and built from asymmetric amino acids, can differentiate between different enantiomers of the same compound when interacting with them on a molecular level. As a result, two enantiomers may have different properties and cause different effects depending on their stereochemistry.

Therefore, the pharmaceutical industry must pay great attention to the medical practices when employing chemical molecules as drugs. For example, ibuprofen is an anti-inflammatory drug (*Figure 2*). (S)-Ibuprofen is significantly more active than (R)-ibuprofen, but commercially the drug is sold as a racemate. However, if enantiomerically pure (S)-ibuprofen is given as drug then a lower dose must be applied as compared to the racemate. Another example is the terpenoid carvone (*Figure 2*); (S)-carvone has a spicy aroma like caraway (and is also a good potato sprouting inhibitor) and restrains the high-fat induced obesity gain. (R)-Carvone, on the other hand, has a sweetish odor like spearmint and is used in air freshening products.
Another infamous example is that of thalidomide (Figure 3) which was primarily prescribed as a sedative, but also to treat morning sickness in pregnant women. \((R)\)-Thalidomide is indeed an effective sedative and easing the sickness, but \((S)\)-thalidomide is severely teratogenic and causes birth defects in fetuses. Ethambutol (Figure 3) is a drug primarily prescribed for treatment of tuberculosis. When the activities of the respective enantiomers were investigated, it was discovered that \((S,S)\)-ethambutol is more potent than \((R,R)\)-ethambutol. Furthermore, it was established that \((R,R)\)-ethambutol causes blindness. The \((R,R)\)-enantiomer was therefore abandoned from use. \((S,S)\)-Ethambutol is instead sold as the pure enantiomer for treatment.

Thereafter, in 1992 the US Food and Drug Administration introduced analysis requirements for both enantiomers of a racemic drug, to minimize potential adverse effects. Out of the top-10 drugs sold in the US, nine are chiral. Therefore, the growing trend and demand for enantiomerically pure compounds is expected to increase research efforts in the field of asymmetric synthesis.

**Enzymes’ Role in Catalysis**

A catalyst is a molecule that increases the rate of a chemical reaction without being consumed in the process. Enzymes are extraordinary catalysts providing tremendous rate enhancements at ambient temperatures and pressure. In 1893, Emil Fischer postulated a “lock and key” hypothesis that describes the enzyme as a lock and the reactant substrate as the key. This theory emphasizes how an enzyme binds and recognizes its substrate but does not explain chemical catalysis. The model was later extended by Linus Pauling in 1948, and he stated that in order to catalyze a reaction an enzyme must be complementary to the *transition state* of the reaction, not the ground state (substrate state). Thus, optimal interactions between enzyme and reactant occur primarily in the transition state.
Thus, a catalyst lowers the activation energy necessary to reach the transition state, and structurally related intermediates of a chemical reaction, as compared to the corresponding uncatalyzed reaction, thereby leading to a rate enhancement of the process (Figure 4). In the process, the enzyme itself often undergoes (a) conformational change(s) when the substrate binds in the active site, due to multiple weak interactions with the substrate. This is referred to as the *induced-fit* postulated in 1958 by Daniel Koshland. The conformational changes can affect a small part of the enzyme active site or can lead to changes in positioning of entire domains. Ultimately, the change brings the required functional groups in the active site into proper positions. Therefore, enzymes may catalyze reactions with high degrees of stereoselectivity. Because of their catalytic power and potential stereoselectivity, the use of enzymes in synthetic chemistry is desirable.

**Figure 4.** An example reaction coordination diagram comparing enzyme catalyzed (black line) and uncatalyzed (gray line) reactions. In the reaction diagram $[E] + [S] \rightarrow [E] + [P]$, the intermediate $[EI]$ occupy minima in the energy progress curve of the enzyme catalyzed reaction compared to uncatalyzed reaction. The $\Delta G^\#_{\text{uncat}}$ and $\Delta G^\#_{\text{cat}}$ corresponds to the activation energy for the uncatalyzed reaction and overall activation energy for the catalyzed reaction, respectively.

**Biocatalysis**

The inclusion of whole cells or isolated enzymes when used in synthetic organic chemistry is known as biocatalysis. This approach has gained momentum during the latter half of the 20th century. Enzymes are biomolecules and are therefore biodegradable, nontoxic and environmentally friendly, in
In comparison to traditional metalloorganic catalysts (Cr, Mn, Ni, Ru, Rh, Pd and Ir). Enzymes facilitate the chemical reactions in aqueous solutions, at moderate values of pH, ambient temperatures and pressure, and are fully renewable. In addition, enzymes being genetically encoded, can be optimized by engineering. Similarly, metabolic pathways can also be engineered and applied as synthetic factories for production of complex organic molecules and thereby contribute to new biocatalytic applications.

Therefore, biocatalytic approaches are attractive to the chemical industry for the possibility of greener manufacturing and new approaches promoting a bio-based economy. For example, it may become possible to utilize biomass from residues of agriculture crops, animals or industrial waste or plain sewage, as new industrial feedstocks for the production of valuable chemicals and fuels. This could also facilitate the transition from a high dependency on finite petroleum and fossil fuel feedstocks. Furthermore, biocatalysis enables in vivo and in vitro technologies that can facilitate new types of chemical synthesis that are otherwise difficult to manufacture by traditional synthetic chemistry.

Kinetic Resolution

Kinetic resolution (KR) is a method used for the separation of enantiomers from a starting racemic mixture employing a stereoselective catalyst. KR mainly relies on the rate difference between the reactions involving two enantiomers, which leads to the enrichment of one enantiomeric product Figure 5.

![Figure 5](image.png)

**Figure 5.** Example of (R) selective kinetic resolution. \( k \) is the rate constant for the reaction of each enantiomer. The catalyst can be an enzyme.

The maximum possible yield via KR is 50%. In other words, an ideal catalyst would transform the starting material to 50% conversion into product, and then stop, resulting in 100% enantiomeric excess (ee) of the (R)-selective product. Enantiomeric purity can be calculated using the following Equation 1 and 2. However, in nature very few enzymes exhibit such excellent enantioselectivity for synthetically relevant chemical transformations. Therefore,
most enzyme reactions will stop at 40 – 45% conversion of reactant to yield a product that is enantiomerically enriched to an acceptable limit. To compare the selectivity for different catalysts, the enantiomeric ratio, ($E$-value) is used. The $E$-value is defined as the ratio between the rate constants of two enantiomers in a catalytic reaction, such as $E = k_{\text{fast}}/k_{\text{slow}}$. For a given reaction, the $E$-value is constant and calculated according to either of the Equations 3 – 5 $^{46-49}$ which are based on the enantiomeric excess of the respective products ($ee_p$), the enantiomeric excess of the reactant substrate ($ee_s$) and the degree of conversion into product ($c$). For example, an enzyme catalyzed reaction one can determine $E$ value by using the ratio of $k_{\text{cat}}/K_M$ for the reactions with the individual enantiomeric substrates. $E = (k_{\text{cat}}/K_M)^R/(k_{\text{cat}}/K_M)^S$. $^{46-49}$

\[
% \text{ee} = \frac{\text{major enantiomer } [R] - \text{minor enantiomer } [S]}{\text{total enantiomers } [R] + [S]} \times 100 \quad \text{Equation 1}
\]

\[
% \text{ee} = \frac{[F_R - F_S]}{100} \quad \text{or} \quad \text{Equation 2}
\]

\[
E = \frac{\ln[1-c(1-ee_s)]}{\ln[1-c(1+ee_s)]} \quad \text{Equation 3}
\]

\[
E = \frac{\ln[1-c(1+ee_p)]}{\ln[1-c(1-ee_p)]} \quad \text{Equation 4}
\]

\[
E = \frac{\ln[ee_p(1-ee_s)/(ee_p+ee_s)]}{\ln[ee_p(1+ee_s)/(ee_p+ee_s)]} \quad \text{Equation 5}
\]

Industrial Biocatalysis and its Applications

In order to become implemented into existing manufacturing processes, enzymes must tolerate the applied reaction conditions such as the presence of organic solvents and/or elevated temperatures and to provide the required substrate scopes, including stereo- and or regioselectivity.

Figure 6. Examples of industrial applications employing biocatalysis.
Therefore, enzymes are often exposed to various degrees of engineering to meet the requirements of the synthetic process. Enzymes may not replace all the present industrial process, but can be used for key steps where high selectivity and purity is a priority. For example, an engineered aminotransferase catalyzed transformation of a prochiral ketone into the chiral amine Sitagliptin, and engineered Baeyer–Villiger Monooxygenase (BVMO) catalyzed sulfide oxidation for Esomeprazole (Figure 6).\(^5^0\)

Over the past decades the increased use and adaptation of enzyme engineering has led to numerous remarkable applications in biocatalysis. Many biochemical transformations produce precursors to final drug molecules or to important building blocks.\(^4^1,^5^1\) The use of engineered ketoreductases has led to synthesis of precursors or intermediates for the following pharmaceutical drugs, montelukast (singulair), crizotinib (xalkori), ezetimibe (zetia, vytorin), atazanavir (reyetaz), atorvastatin (lipitor), and also to generic drugs (\(R\))-phenylephrine and (1\(S\),2\(S\))-pseudoephedrine were developed by Codexis.\(^4^1\)

Traditional Limitations of Enzymes as Biocatalysts

Despite many success stories, in order to become the industrial catalysts of choice in synthetic chemistry, enzymes still need further optimizations. The listed drawbacks (Table 4) not only limit the enzyme applications in a practical way, but also hinders breakthroughs of biocatalysis in general.\(^4^0,^4^8\)

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemo-, regio-, and enantioselectivity</td>
<td>Insufficient selectivity and narrow substrate scopes</td>
</tr>
<tr>
<td>Active in aqueous solutions under mild conditions</td>
<td>Insufficient stability in non-aqueous solvents, pH or at extreme temperature</td>
</tr>
<tr>
<td>Biodegradable after use</td>
<td>Inhibition (substrate/product)</td>
</tr>
<tr>
<td>Renewable (recovery and recycling)</td>
<td>Cofactor (cost)</td>
</tr>
<tr>
<td>Minimal waste production</td>
<td></td>
</tr>
<tr>
<td>No derivatization</td>
<td></td>
</tr>
</tbody>
</table>

Nonetheless, the desire to use enzymes as biocatalysts has led to setting-up new frontiers for enzymes engineering for specific functions. For example, dedicated databases, metagenomics searches, \textit{de novo} protein design and \textit{in vitro} evolution of enzymes.

Enzyme Engineering

Currently, there are several techniques available for engineering of enzymes. Ranging from totally randomized methods which mimic the genetic variability
that is a component of natural evolution, to sophisticated methods based on qualitative structural information and rational strategies. Most approaches depend on the polymerase chain reaction for introduction of mutations.52,53

**Directed Evolution**

In 1859, Charles Darwin developed the *Evolution Theory*.54 It describes that genetic variation is generated by hereditary recombination between individuals or by introduction of mutations. If the genetic variability is coupled to a selection pressure this ultimately leads to evolution of traits. Directed evolution of proteins conducted in laboratory mimics natural evolution by introducing mutations *in vitro* and defines selection criteria for a desired function.55,56 Improved variants are used as parents for another round, until beneficial mutations have accumulated or when no further improvements is achieved. In general, important traits in enzyme engineering which are target for the subjected changes are substrate scope, selectivity, thermostability and pH optimum. The two most commonly applied mutagenesis procedures, employ error-prone polymerase chain reaction (epPCR)57 or DNA shuffling58 (*Table 5*).

<table>
<thead>
<tr>
<th>Method</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error-Prone PCR (epPCR)</td>
<td>Random mutagenesis</td>
</tr>
<tr>
<td>DNA shuffling</td>
<td>Random mutagenesis</td>
</tr>
<tr>
<td>Iterative Saturation Mutagenesis (ISM)</td>
<td>Targeted/predefined regions</td>
</tr>
</tbody>
</table>

Firstly, epPCR depends on a DNA polymerase, that lacks proof-reading activity, e.g. *Taq* DNA polymerase, and thus leads to incorporation of mutants at relatively controlled frequency, ~ one error per $10^4$ bp, and hence, mimics imperfect gene replication.59,60 Secondly, DNA shuffling is homologous recombination of pools of related genes, that following random fragmentation and re-assembly at cross-over points by PCR-imitates the natural recombination process.58 Furthermore, epPCR and DNA shuffling methods are helpful when structural information or homology models of the protein of interest are not available. Directed evolution have been applied to improve several properties, such as stability, selectivity, substrate scope, and solvent tolerance.59–63 However, both techniques generally generate (very) large gene libraries, which require intensive high-throughput screening efforts in the subsequent searches for improved gene products. Therefore, efforts have been focused to design approaches that combine the features of rational design and random mutagenesis to enable generation of smaller, yet high quality libraries that can be assessed more rapidly.40
Rational Design

In 1978 Michael Smith introduced site-directed mutagenesis (SDM). Using specially designed DNA oligonucleotides as primers for DNA synthesis, one can replace a single amino acid codon at a predetermined site in the corresponding gene/cDNA and thereby change the encoded amino acid in the translated protein by any of the other 19 canonical amino acids.64

Figure 7. Example of site directed mutagenesis (SDM) exchange of base in the gene sequence to create another genotype.

SDM targeted residues are mutated to specific amino acids based on protein structure information (Figure 7). This methodology has paved the way for altering the structure site-specifically and can thereby be applied to affect catalytic properties of enzymes. However, to be able to improve substrate selectivity is often a more challenging task that require a number of mutations.

Semi-Rational Design/Iterative Saturation Mutagenesis

Semi-rational design combines random and rational mutagenesis strategies. This approach, requires structure information for identification of the key amino acid residues that are deemed to impact an enzymes’ catalytic properties, such as catalytic efficiencies, regio-, stereoselectivity, substrate scope or stability. Selected residue positions are randomly mutated by saturation mutagenesis. Typically, two or three amino acid residues located in proximity to each other in the structure are grouped together as one site, and targeted for saturation mutagenesis. If the objective is to evolve an enzyme for altered selectivity and substrate scope, the amino acid residues which are located within the active site are in focus. This methodology is defined as Combinatorial Active-Site Saturation Test (CAST), and was first introduced by Manfred Reetz in 2005.65

The CASTing approach can also be used to improve on thermostability of enzymes by targeting the surface residues which display (constant motion due to thermal and kinetic energy of the atoms) comparably high B-factor values, this is known as B-fit. If the CASTing and B-fit are combined with iterative saturation mutagenesis (ISM) one can generate focused and high-quality
Applying the ISM methodology, one can perform iterative cycles of saturation mutagenesis and stepwise visit different mutation sites in the target structure. A hit from the first-round library and selection becomes the parent for the next cycle (Figure 8). This process is repeated until satisfactory improvements have been obtained.

If different sites are visited, ISM might allow for the phenomenon of epistasis in which the effect of one mutation is dependent on the presence of another mutation. This is in contrast to the additive effect of combinations of individual modifications. In short, the ISM method includes: i) identification and ranking of amino acid residues, ii) prioritizing and grouping of target amino acid residues, and finally iii) a screening effort. The ranking of amino acid residues in the process is important, since prioritization is key to keep the number of possible variants and screening efforts in manageable limits for the experimentalist. Typically, libraries are constructed with two or three amino acid residues. To further minimize the library size, degenerated codon sets can be used. A common alternative is an ‘NDT’ mixture of codons, that encodes 12 codons of the 20 natural amino acids, without codon bias (Table 6).

Table 6. Amino acids involve in NDT codon degeneracy.

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>NDT*</th>
<th>Side chain Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>F L I V Y</td>
<td>Hydrophobic</td>
<td></td>
</tr>
<tr>
<td>H R D N S C</td>
<td>Charged/polar</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Small</td>
<td></td>
</tr>
</tbody>
</table>

* = all bases, D = A/G/T and T = T
NNK or NNS (K = G or T, S = G or C) codon sets, on the other hand allows the all 20 amino acids and also one stop codon, but involves the 32 possible codons. Therefore, several amino acids will be over represented in the library leading to unwanted bias. For example, a site composed of two amino acids, NNK requires almost 3066 clones, due to necessary oversampling, whereas NDT requires only 430 clones to screen the necessary 95% coverage of relevant protein sequence (Table 7).

Table 7. NNK and NDT degeneracy codon sets for necessary 95% coverage.

<table>
<thead>
<tr>
<th>Amino acid No*</th>
<th>NNK Codons</th>
<th>NNK Screening Variants</th>
<th>NDT Codons</th>
<th>NDT Screening Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>94</td>
<td>12</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>1028</td>
<td>3066</td>
<td>144</td>
<td>430</td>
</tr>
<tr>
<td>3</td>
<td>32768</td>
<td>98163</td>
<td>1728</td>
<td>5175</td>
</tr>
<tr>
<td>4</td>
<td>1048576</td>
<td>3141251</td>
<td>20736</td>
<td>62118</td>
</tr>
<tr>
<td>5</td>
<td>33554432</td>
<td>100520093</td>
<td>248832</td>
<td>745433</td>
</tr>
<tr>
<td>6</td>
<td>&gt; 1.0 x 10⁹</td>
<td>&gt; 3.2 x 10⁹</td>
<td>&gt; 2.9 x 10⁹</td>
<td>&gt; 8.9 x 10⁹</td>
</tr>
<tr>
<td>7</td>
<td>&gt; 3.4 x 10¹⁰</td>
<td>&gt; 1.0 x 10¹¹</td>
<td>&gt; 3.5 x 10⁷</td>
<td>&gt; 1.1 x 10⁸</td>
</tr>
<tr>
<td>8</td>
<td>&gt; 1.0 x 10¹²</td>
<td>&gt; 3.3 x 10¹²</td>
<td>&gt; 4.2 x 10⁸</td>
<td>&gt; 1.3 x 10⁹</td>
</tr>
<tr>
<td>9</td>
<td>&gt; 3.5 x 10¹³</td>
<td>&gt; 1.0 x 10¹⁴</td>
<td>&gt; 5.1 x 10⁹</td>
<td>&gt; 1.5 x 10¹⁰</td>
</tr>
<tr>
<td>10</td>
<td>&gt; 1.1 x 10¹⁵</td>
<td>&gt; 3.4 x 10¹⁵</td>
<td>&gt; 6.1 x 10¹⁰</td>
<td>&gt; 1.9 x 10¹¹</td>
</tr>
</tbody>
</table>

* Number of amino acid positions at one site

To overcome the problems of biased codon degeneracy, one can construct “intelligent” libraries. By using a combination of NDT – 12 codons, VMA – 6 codons (E, A, Q, P, K, and T), ATG and TGG codes sets, all 20 amino acids without bias are encoded. Overall, the decision about ‘how to restrict the library size’ often depends on deciding factors, such as screening methodology and available structural knowledge of the targeted enzyme.

Screening and Selection

Despite reduction of library size to become manageable for efficient screening of an adequate number of variants, the screening step is often a bottleneck in enzyme engineering. If more than two or three residues are modified at the same time, the number of transformants rapidly increases (Table 7). Therefore, one strives to use medium- to high-throughput techniques for screening and selection. Apart from the library size and the efforts of the experimenter, other influencing factors are efficiency of screening assays. For example, i) can the method be generally applicable for different study cases, ii) robustness: results are trustworthy and reproducible, iii) operative costs and efficiency: reagents and equipment. Generated library samples, often in the form of bacterial cell lysates are assayed in flat-bottom 96-well microtiterplates and
then screened for the feature of interest e.g. catalytic activity. For example, detection of active variants of alcohol dehydrogenase can exploit the differences in the spectral properties of the reduced and oxidized cofactor NADH/NAD$^+$ for direct real-time measurements in a spectrophotometer (Figure 9).$^{69,70}$

**Figure 9.** Example (A) screening of alcohol dehydrogenase transformants for conversion of secondary alcohol into corresponding ketone, (B) coupled screening assay of retro-aldol cleavage activity of D-fructose 6-phosphate aldolase transformants utilizing an (engineered) aldehyde reductase (FucO DA1472) as reporter enzyme.$^{69}$

Using the described screening method, one can measure activities of a whole enzyme library and from generated data identify the best variants. Potential hits are validated by further re-screens, gene sequencing and detailed functional analysis. If an enzyme is not dependent on a cofactor or not possible to measure quantitatively by other methods and high-throughput techniques, then one can attempt to design coupled assays involving reporter enzymes that can provide a spectroscopic signal$^{69,70}$ (Figure 9B), using end-point colorimetric assays$^{71}$, product analysis by gas or liquid chromatography, cell sorting (FACS) or micro-fluidic devices.$^{72,73}$ Furthermore, several time-consuming and repetitive tasks such as picking bacterial colonies, microbe cultivation and libraries screenings can be automated for even higher-throughput.

Characterization

Isolated variants need to be characterized and validated. Often, characterization is time consuming, therefore only the most promising variants will undergo detailed analysis of kinetics, selectivity, or lack thereof (promiscuity), and stability.

Enzyme Kinetics

The catalytic efficiency of a purified enzymes can be determined by measuring the velocity/rate of a catalyzed reaction. Experimentally, one can measure
on different time-scales following the reaction from seconds to minutes during the steady-state reaction phase or shorter, milliseconds to seconds during the pre-steady-state reaction phase.

**Steady-State Kinetics**

In steady-state kinetic experiments initial rates/velocities $v_0$ are measured. For simplification the method involves two the assumptions that $\Delta[ES]$, and $\Delta[S]$ and $[P]$ are considered to be negligible.

\[
E + S \xrightarrow{k_1} ES \xrightarrow{k_2} E + P
\]

*Figure 10.* A simplest example of an enzyme kinetic mechanism where the substrate and enzyme form a Michaelis-Menten complex ES, subsequently breaks down either forward product into free E and P or back into again free E and S.

For example, a simple two-step reaction shown in (Figure 10)\(^{16}\) was first postulated in 1913 by Leonor Michaelis and Maud Menten.\(^{74-76}\) At relatively low concentrations of substrate, $v_0$ increases almost linearly with an increase in $[S]$, and at higher substrate concentrations $v_0$ increases by smaller and smaller amounts, and to a point where an increase in $[S]$ has no influence on $v_0$. Reactions where the catalyzing enzyme is saturated at this plateau-like region are close to the maximum velocity, $V_{\text{max}}$. The relation between the reaction velocity $v_0$ and concentration of substrate for a given concentration of enzyme can be visualized in a saturation curve (Figure 11).

\[v_0 = \frac{V_{\text{max}}[S]}{K_M} \]

*Figure 11.* Steady-state kinetics, a hyperbolic curve (solid gray line) shows typical behavior of “initial velocities”, $v_0$, for varied high and low substrate concentrations ([S]) for a constant ([E]) concentration. This type of behavior described by the Michaelis-Menten equation (Equation 4). The $V_{\text{max}}$, maximum velocity (dotted gray line), and the corresponding ([S]) concentration at $1/V_{\text{max}}$ (dotted light gray line) represents the Michaelis-Menten constant, $K_M$. $V_{\text{max}}$ is determined by curve fitting [S] and from this value the turnover number $k_{\text{cat}}$, and specificity constant $k_{\text{cat}}/K_M$, can be obtained.
Typically, most enzymes that exhibit a hyperbolic dependence of $v_0$ on $[S]$ obey Michaelis-Menten kinetics (Equation 6), which is useful in practical determinations of $K_M$, $V_{max}$ and $k_{cat}$.\(^{16}\)

$$v_0 = \frac{V_{max} [S]}{K_M + [S]} \quad \text{Equation 6}$$

From the Michaelis-Menten equation we can derive three constants $K_M$, $k_{cat}$ and $k_{cat}/K_M$. The term $K_M$ is Michaelis constant equals the $[S]$ where $v_0$ is at $\frac{1}{2}V_{max}$. Therefore, both $K_M$ and $V_{max}$ are substrate dependent, when $k_2$ is rate-limiting then $K_M$ approximates to $k_1/k_{-1}$ which defines the dissociation constant $K_d$ of enzyme-substrate complex. Under such conditions $K_M$ can provide a measure of affinity of the enzyme for its substrate in $[ES]$ complex. However, this scenario does not apply to many enzymes and sometimes $k_2 > k_{-1}$ then $K_M$ is defined by $(k_2 + k_{-1}/k_1)$ and $K_M$ remains a complex function of all three rate constants for formation and decay of ES (Equation 7).\(^{16}\)

$$K_M = \frac{k_{-1} + k_2}{k_1} \quad \text{if} \quad k_2 \ll k_{-1}, \quad K_M \approx \frac{k_{-1}}{k_1} \approx K_d \quad \text{Equation 7}$$

Figure 12. Example, the kinetic mechanism constitutes an additional intermediate Enzyme-product formation, and release of free enzyme and product.

In order to compare the different enzyme-catalyzed reactions at saturation, a more general rate constant, $k_{cat}$, the turnover number is introduced. If the catalyzed reaction consists of more than step following formation of enzyme-substrate complex, then $k_{cat}$ becomes a composite of all rate constants leading to free enzyme and product (Figure 11). In the example shown in Figure 12, either $k_2$ or $k_3$ may contribute to the value of $k_{cat}$ (Equation 8).

$$v_0 = \frac{k_{cat} [E]_t [S]}{K_M + [S]} \quad \text{Equation 8}$$

Using the kinetic parameters $k_{cat}$ and $K_M$ we can compare the different enzymes catalytic efficiencies $k_{cat}/K_M$. If the $v_0$ is dependent on $[E]_t$ and $[S]$ then the reaction rate is second-order and $k_{cat}/K_M$ is a second-order rate constant (Equation 9).\(^{16}\)

$$v_0 = \frac{k_{cat}}{K_M} [E]_t [S] \quad \text{Equation 9}$$

Enzymes may undergo conformational changes upon substrate binding as part of their mechanism which can in turn affect the affinity for other substrates or
control the overall reaction. Such cases can be determined from the possible presence of cooperative behavior in substrate concentration dependence. Cooperative enzymes do not follow the Michaelis-Menten kinetics, strictly. Positive and negative cooperativity can be determined by fitting a modified Hill equation (Equation 10) to the steady-state kinetic data points.

\[
\frac{v_0}{[E]_t} = \frac{k_{\text{cat}} [S]^n}{K_{0.5}^n + [S]^n}
\]

Equation 10

Figure 13. Saturation curves showing different types of cooperativity. The gray line represents the hyperbolic curve as defined by the Michaelis-Menten equation is \(n=1\). The dotted gray line represents negative cooperativity with a Hill-coefficient is \(n<1\), while the positive cooperativity and black curve with a Hill-coefficient is \(n>1\).

In the above equation \(K_{0.5}\) replaces \(K_M\). The Hill-coefficient \(n\), describes the level of cooperativity. If the observed Hill-coefficient \(>1\) the enzyme displays positive cooperativity, and if \(<1\), is negative cooperativity. Typical saturation curves are shown in Figure 13.

**Pre-Steady-State Kinetics**

The steady-state kinetics only provides information about the overall reaction, but does not provide individual microscopic rates on the reaction pathway. To get more complete information generally requires methods that allow for measurements in the time period leading up to the steady state. Therefore, pre-steady-state kinetic experiments can be achieved with e.g. stopped-flow methods. Using the stopped-flow spectro/fluorophotometer one can identify rate-limiting and rate-determining steps in a reaction (Figure 14), such as release of cofactor or chemical transformations.

For example, NADH binding to the enzyme can be detected by monitoring the exponential burst increase in NADH fluorescence upon binding of the coenzyme to the active site, and likewise, NAD\(^+\) binding to the enzyme can be detected by monitoring the quenching of intrinsic tryptophan or tyrosine
fluorescence, respectively, for detection of binding and dissociation rates. To determine observed rates $k_{\text{obs}}$, the progression curves are fitted to Equation 11, a function of a single exponential, with a floating end point. ‘$F$’ is the recorded spectroscopic signal, ‘$A$’ the signal amplitude, ‘$t$’ is time and ‘$c$’ is the floating endpoint.

$$F = A \exp(-k_{\text{obs}} t) + C$$  \hspace{1cm} \text{Equation 11}$$

The kinetic parameters can subsequently be determined by fitting the substrate/ligand concentration dependence of $k_{\text{obs}}$ to a secondary model. For instance, rates of binding and dissociation of NAD$^+$ or NADH can be determined by fitting a linear function (Equation 12) to the determined $k_{\text{obs}}$ values, assuming a 1:1 binding event.

$$k_{\text{obs}} = k_{\text{on}} [\text{NAD}(\text{H})] + k_{\text{off}}$$ \hspace{1cm} \text{Equation 12}$$

In the case of determining rates of chemical steps, e.g. alcohol dehydrogenase catalyzed oxidation of an alcohol into the corresponding carbonyl compound, the observed rates determined by fitting Equation 11 to the experimental progression curves (i.e. followed as the formation of NADH upon reduction of NAD$^+$) are in turn fitted to another secondary equation (Equation 13). From the substrate dependence of $k_{\text{obs}}$, the kinetic parameters for oxidation, $k_{\text{ox}}$ and $K_S$ (the dissociation constant of the enzyme-alcohol complex) can be determined. For Equation 13 to be a valid model, it is required that a reaction step downstream of the analyzed step (in this example alcohol oxidation), to be slower than the analyzed step. If this is not the case, $k_{\text{obs}}$ will equal the steady state rate ($v_0/[E]_0$).

$$k_{\text{obs}} = \frac{k_{\text{ox}} [S]}{K_S + [S]}$$  \hspace{1cm} \text{Equation 13}$$

*Figure 14. Example of a pre-steady-state exponential phase reaction burst ($k_{\text{obs}}$) and followed by steady-state phase.*
Enzyme Substrate Specificity and Selectivity

Enzymes can be exquisitely specific and selective for their native substrates, although often lacks adequate activities with industrially relevant reactants. Enzyme selectivity arises from their three-dimensional structure of the active site, which is complementary to the transition state of the reaction. Therefore, enzyme selectivity towards different substrates can, to a degree, be simplified by complementary interactions between active site residues and the respective substrates. It follows that the active site may stabilize the transition states for reactions involving different substrates which varying degree of efficiencies leading to differences in catalytic activity (Figure 4).

Figure 15. Different types of enzyme selectivity, (A) substrate selectivity, (B) chemoselectivity, (C) regioselectivity, and (D) enantioselectivity. Enzymes may display several of the following types of selectivity simultaneously.

Enzyme specificity/selectivity between different substrates can be measured by comparing the $k_{cat}/K_M$ values. Enzyme selectivity can refer, to substrate-, chemo-, regio- and/or enantioselectivity (Figure 15).

Nonproductive binding

In an enzymatic reaction a substrate may bind in a nonproductive (unreactive) mode that may result in competition with the productive (reactive) binding mode (Figure 16). For example, anilide substrates of chymotrypsin competitively bind in a nonproductive mode with the anilide ring in the substrate binding pocket compared to the corresponding acetyl and trimethylammonium substituted N-acetyl-L-tyrosineanilides.
Nonproductive binding of substrate effectively lowers the steady-state constants $k_{\text{cat}}$ and $K_M$. Using the following equations 14 – 16 we can determine the nonproductive binding.\(^\text{78,81}\) Hence, the value of $k_{\text{cat}}/K_M$ is unchanged.

\[
\nu_0 = \frac{[E]_t [S] k_2}{K_S + [S] (1 + K_S/K'_S)} \quad \text{Equation 14}
\]

\[
k_{\text{cat}} = \frac{k_2}{1 + K_S/K'_S} \quad \text{Equation 15}
\]

\[
K_M = \frac{K_S}{1 + K_S/K'_S} \quad \text{Equation 16}
\]

**Promiscuity**

As described above, enzyme selectivity in substrate recognition is thought to be essential for catalysis. However, many enzymes are capable of catalyzing multiple reactions that transform non-physiological substrates into products.\(^\text{82–84}\) This (lack of) selectivity is referred to as enzyme promiscuity. Enzyme promiscuity has received considerable attention to highlight the potential mechanistic and evolutionary implications for enzyme-substrate recognition and as a model for *in vitro* manipulations with the aim to alter an enzyme’s function for a better fit to an industrial applications.\(^\text{85–89}\) Moreover, the relaxed promiscuous nature of the enzymes is not a new discovery,\(^\text{90}\) earlier examples of enzyme promiscuity includes carbonic anhydrases,\(^\text{91}\) pepsin,\(^\text{92}\) chymotrypsin,\(^\text{93}\) L-asparaginases\(^\text{94}\) and pyruvate decarboxylases.\(^\text{84}\) More recent experimental evidences indicate that promiscuity is not only in interactions but also in the native functions of proteins. Thus, promiscuity can be important in the evolution of secondary activities in organisms, and may provide increased abilities to adjust to changes in the environment. Presumably, enzyme promiscuity have evolved from gene duplications and resulting in the incorporation of the respective activities under a new selection pressure.\(^\text{82–84}\)

Promiscuity can be grouped into different classes, \(i\) substrate promiscuity – functional groups, regio- and enantioselective, \(ii\) catalytic promiscuity – multiple reactions, single reaction types with different substrates, \(iii\) conditions promiscuity – anhydrous solvents, temperature and pH, \(iv\) product promiscuity – generating different products from the same reaction, \(v\) causes of
promiscuity – multiple domains are used to enable single proteins to achieve multiple functions, conformational flexibility and dynamics, and vi) protein folding promiscuity – architecture, plasticity and ability to present catalytic amino acid residues surrounding the active site.\textsuperscript{82–84} Therefore, understanding protein promiscuity is becoming increasingly important not only to optimize enzyme engineering applications for industry and academia, but may also be of increasing importance in areas of synthetic biology, metagenomics and drug development.

Structure-Function Relationship

To know the tertiary structures of proteins in detail is highly valuable in enzyme engineering efforts. By accessing protein structures, one can assign putative roles of catalytic amino acids residues to catalysis, or to binding interactions with substrates, inhibitors and cofactors. Knowledge of the tertiary structure is also crucial in studies of protein folding/unfolding and molecular dynamics. A protein structure can be determined by X-ray crystallography,\textsuperscript{95} nuclear magnetic resonance spectroscopy (NMR)\textsuperscript{96} or cryo-electron microscopy (cryo-EM).\textsuperscript{97} Using the three-dimensional structures of protein, one can construct computational methodology to understand the flexibility, dynamics, force fields, folding and conformational energy landscape to employ the design or engineer for enzyme catalysis.\textsuperscript{98}
Alcohol Dehydrogenases

Enzymes that catalyze chemical redox reactions are classified as oxidoreductases, and can be found in all organisms. Oxidoreductases are divided into subclasses depending on electron donor and electron acceptor. Common cofactors are flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), nicotinamide adenine dinucleotide (NAD+/NADH), and nicotinamide adenine dinucleotide phosphate (NADP+/NADPH). A subclass of oxidoreductases, alcohol oxidoreductases (EC 1.1.X.X), interconvert alcohols to the corresponding aldehydes or ketones in the presence of cofactor to synthesize metabolites in biological system.

Alcohol dehydrogenases (ADH) can be favorably employed in oxidation and reduction reactions for organic synthetic applications, e.g. fermentation of glucose into ethanol or for asymmetric synthesis of chiral alcohols. For example, in the last decades many alcohol dehydrogenases have been isolated from different organisms ranging from S. cerevisiae, horse liver, Thermoanaerobium brockii, Rhodococcus erythropolis, Rhodococcus ruber, Candida parapsilosis, Lactobacillus brevis, Lactobacillus kefir, Pseudomonas sp.

This thesis is focused on alcohol dehydrogenases which are depending on NAD+/NADH as a cofactor (EC 1.1.1.1) in redox chemical transformations. In general, alcohol dehydrogenases catalyze the reaction through hydride transfers between NAD+/NADH and the alcohol or the corresponding carbonyl compound.

![Figure 17. Example of alcohol dehydrogenase catalyzed ketone reduction that obeys the Prelog rules.](image)

Although, these enzymes are further sub-divided into three classes; i) class-I: medium length (327 – 376 amino acid residues per chain) zinc dependent, ii)
class-II: short length (approximately 250 amino acid residues per chain) zinc dependent and class-III: iron activated (approximately 385 amino acid residues per chain) alcohol dehydrogenases. Many of the known alcohol dehydrogenases are (S)-selective and obey the so-called Prelog rule (Figure 17). This rule dictates that hydride transfer preferentially occurs to the re-face of a prochiral ketone to lead the corresponding (S)-configured alcohols. However, there are some alcohol dehydrogenases, such as Pseudomonas sp.-ADH that display reversed stereopreference and deliver the hydride to the si-face of a prochiral ketone leading to the corresponding (R)-alcohols (Table 8).

Table 8. Examples of Prelog and anti-Prelog alcohol dehydrogenases.

<table>
<thead>
<tr>
<th>Alcohol dehydrogenase</th>
<th>Cofactor</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. cerevisiae “baker’s yeast”-ADH</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Horse liver-ADH</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Rhodococcus ruber-ADH-A</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Rhodococcus erythropolis-ADH</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Candida parapsilosis-ADH</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Hydroxysteroid-ADH</td>
<td>NADH</td>
<td>Prelog</td>
</tr>
<tr>
<td>Thermoanaerobium brockii-ADH</td>
<td>NADPH</td>
<td>Prelog*</td>
</tr>
<tr>
<td>Lactobacillus brevis-ADH</td>
<td>NADPH</td>
<td>anti-Prelog</td>
</tr>
<tr>
<td>Lactobacillus kefir-ADH</td>
<td>NADPH</td>
<td>anti-Prelog</td>
</tr>
<tr>
<td>Pseudomonas sp.-ADH</td>
<td>NADH</td>
<td>anti-Prelog</td>
</tr>
</tbody>
</table>

*anti-Prelog specificity when subjected to smaller ketones.

Structural Similarities of ADHs and Reaction

The comparison of alcohol dehydrogenases structural features provides information about evolutionary and structure-function relationships. Alcohol dehydrogenases that are found in higher eukaryotes, such as plants and animals, are usually dimeric, whereas in lower eukaryotes (yeast) and prokaryotes the enzymes are tetrameric. Among the alcohol dehydrogenases in this superfamily, the horse liver ADH (homodimeric, with 374 amino acids per subunit) and the ADH from S. cerevisiae (homotetrameric, with 347 amino acid per subunit) have been studied most extensively. Although the sequence identity of the horse-liver and yeast enzymes is only ~24%, molecular modeling revealed that the three-dimensional structures are very similar. In both structures, each subunit contains two domains. Firstly, a coenzyme binding domain of the Rossmann fold (six-stranded parallel β-pleated sheet with two helices on each side of the sheet), and secondly, a substrate binding/catalytic domain that contains one catalytic Zn²⁺ ion to which the substrate coordinates and a structural zinc ion in a distant loop.

The substrate (alcohol and ketone) binds in a cleft between the two domains and binding can vary, reflecting the substrates which a particular enzyme acts
upon, e.g. ADHs from horse-liver and *S. cerevisiae*. Both of these ADHs belong to the same class I alcohol dehydrogenases and their catalytic cycles follows an ordered bi-bi mechanisms. Both enzymes have been proposed to exists in either a “closed” or “open” conformation, that depend on, and influences the cofactor binding and release. The oxidation of an alcohol begins with coenzyme (NAD\(^+\)) binding and ends with NADH release. The hydride transfer from substrate to NAD\(^+\) is facilitated by the catalytic zinc that acts as a Lewis acid to lower the pK\(_a\) of the coordinated alcohol. This leads to dissociation of alcohol proton through a proposed proton transfer chain involving enzyme side chains and the 2’-hydroxyl of bound nucleotide ribose (*Figure 18*A). The formed alcoholate is stabilized by the catalytic Zn\(^{2+}\). Thereafter, hydride transferred from the \(\alpha\)-carbon of the alcoholate to the C-4 of NAD\(^+\) (*Figure 18*B).

**Figure 18.** Schematic illustration of the proton and hydride transfer mechanism proposed to occur in horse liver ADH\(^{107,109}\): (A) The alcohol proton is transferred to the enzyme side chain of Ser\(_{48}\), and further shuttled via Grothrus mechanism to the nucleotide ribose and His\(_{51}\) and delivery to bulky water (B) Hydride transfer from the alcoholate ion to C-4 carbon of the coenzyme NAD\(^+\).

**Alcohol Dehydrogenase from *Rhodococcus ruber***

The investigations in this thesis work have been primarily focused on ADH-A from *Rhodococcus ruber* DSM 44541,\(^{112,113}\) which belongs to the class I alcohol dehydrogenase and depend on the same cofactors as described for the horse-liver enzyme – two Zn\(^{2+}\)ions and the organic coenzyme NAD\(^+\)/NADH. The tertiary structure of ADH-A has been determined at 2 Å resolution (pdb ID: 3jv7)\(^{114}\) with four polypeptide chains (homotetramer) in asymmetric unit (*Figure 19*). Although the protein appears to be tetramer, in solution the functional form is a homodimer. Each subunit consists of 345 amino acids,\(^{114}\) with the catalytic zinc ion is surrounded by C38, S40, H62 and N153. The active
site entrance is partially formed by a loop from the neighboring subunit of the dimer involving mainly three phenylalanine residues F281, F282 and F286.\textsuperscript{114}

The structure of wild type ADH-A is highly similar to the horse-liver ADH (pdb ID: 1HLD). Due to the structural similarities between ADH-A\textsuperscript{114} and horse-liver ADH\textsuperscript{107}, it can be assumed that the catalytic mechanisms are highly similar (Figure 18).

Figure 19. The tertiary structure of one subunit of alcohol dehydrogenase A from \textit{R. ruber} (gray). The cofactor, NAD\textsuperscript{+} shown in stick representation (dark gray) and the catalytic Zn\textsuperscript{2+} is shown as a black sphere. The structural Zn\textsuperscript{2+} is away from the active site. Image was constructed in PyMol (www.pymol.org), pdb ID: 3jv7.

A particularly interesting feature of ADH-A is the catalyzed interconversion of sec-alcohols and the corresponding ketones, with high enantio- and regioselectivity. The enzyme is also unusually resistant to water miscible organic solvents, such as 50\% (v/v) acetone, 80\% (v/v) isopropanol and has been demonstrated to be active even in 99\% (v/v) hexane.\textsuperscript{112,113,115,116} ADH-A accepts a relatively wide range of aromatic sec-alcohols as substrates with a preference for (S)-enantiomer.\textsuperscript{112–119} The above-mentioned traits are facilitating ADH-A as a viable enzyme to study and engineer for redox biocatalytic transformations for industrial setting. To overcome drawback of the ADH-A cofactor dependency which introduce unacceptable high costs for industrial
processes, it is possible to integrate processes that regenerate the cofactor by coupling with either a coupled-substrates (Figure 20) or by coupled-enzymes approaches.112–118

Figure 20. An example of in situ cofactor regeneration reaction by ADH in a coupled-substrate approach. Oxidation of a sec-alcohol substrate is followed by re-oxidation of NAD(P)H via reduction of acetone to isopropanol.
The present investigation has been primarily focused on engineering of ADH-A and characterization of resulting enzyme variants as potential biocatalysts for chemical transformations. As described in the earlier section, ADH-A tolerates water miscible organic solvents, accepts a relatively wide range of aromatic sec-alcohols/ketones as substrates in redox reactions. These characteristics caught our interest to investigate if ADH-A could be evolved in vitro to acquire new predesigned properties. The engineering efforts have aimed for altered substrate scope, including altered stereo- and regioselectivities. Furthermore, possible substrate promiscuity in engineered enzyme variants as compared to the wild type parent enzyme, was also studied.

Paper I

Wild type ADH-A exhibits a 270-fold preference towards the (S)-enantiomer of 1-phenylethanol over the corresponding (R)-enantiomer. The lower catalytic efficiency with (R)-1-phenylethanol was attributed due to the nonproductive binding of this enantiomer in the active site.119

Library Generation

To investigate the possibility to switch the enantioselectivity of the wild type into instead favoring the (R)-enantiomer, and to also study the possible involvements of nonproductive substrate binding as a mechanism in substrate discrimination, we subjected the ADH-A active site to iterative saturation mutagenesis. Based on the available crystal structure information, we targeted three subsites (A, B and C), that contain residues contributing to the active site cavity. At each mutation site, two amino acid residues placed spatially adjacent to each other were grouped together. The chosen subsites are located on opposite sides of the active site cavity; The A site contains Y294 and W295, the B site, Y54 and L119, and the C site, F43 and I271, respectively (Figure 21).120 The variant enzymes encoded in the respective libraries were co-expressed with chaperonins GroEL/ES to assist in their folding to native, soluble protein structures.121
Evolutionary Routes
In site “A”, W295 is involved in interactions with P89 through Van der Waals (VdW) contacts, to T152 through hydrogen bonding between the indole nitrogen and the backbone carbonyl oxygen and T157 and is in VdW contact with the side chain. Site “B” involves Y54 and L119 and site C, F43 and I271 where the side chain of the latter is in VdW contact with the nicotinamide ring of the coenzyme. The gene libraries were constructed by PCR with a combination of degenerate codons (NDT, 12 codons, VMA – 6 codons E, A, Q, P, K, and T, ATG and TGG) to access all 20 amino acids without bias.68 The three targeted sites were exposed to saturation mutagenesis in a stepwise manner (Figure 8: Iterative Saturation mutagenesis).65–67 Firstly, the “A” site was screened and two isolated ‘hits’ (A1 and A2) that displayed apparently improved activity with (R)-1-phenylethanol were characterized. Subsequently, A2 was chosen to parent a second generation of ISM that targeted the “C” site. From this A2C library, three variants were isolated and characterized (A2C1, A2C2 and A2C3). Thereafter, A2C2 was chosen as parent for the third generation of
ISM, moving to the “B” site, which provided one variant with further improved activity A2C2B1.

Table 9. Selected enzyme variants from different generations and their mutations.

<table>
<thead>
<tr>
<th>Variants</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Other*</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>Y294</td>
<td>W295</td>
<td>Y54</td>
<td>L119</td>
</tr>
<tr>
<td>A1</td>
<td>°</td>
<td>A</td>
<td>°</td>
<td>°</td>
</tr>
<tr>
<td>A2</td>
<td>F</td>
<td>A</td>
<td>°</td>
<td>°</td>
</tr>
<tr>
<td>A2C1</td>
<td>F</td>
<td>A</td>
<td>°</td>
<td>°</td>
</tr>
<tr>
<td>A2C2</td>
<td>F</td>
<td>A</td>
<td>°</td>
<td>°</td>
</tr>
<tr>
<td>A2C3</td>
<td>F</td>
<td>A</td>
<td>°</td>
<td>°</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>F</td>
<td>A</td>
<td>F</td>
<td>°</td>
</tr>
<tr>
<td>A2C2B2</td>
<td>F</td>
<td>A</td>
<td>W</td>
<td>°</td>
</tr>
</tbody>
</table>

*Mutations was introduced from PCR artifact

Isolated variants (Table 9) were studied for their catalytic activities with (R)-1-phenylethanol, (S)-1-phenylethanol, acetophenone and 2-propanol, respectively. The steady-state kinetics were conducted to determine the effect of the different mutations. In addition, the activities with deuterated 2-propanol-2-d1, were analyzed in order to map rate-limiting steps of the reaction. Additionally, applying stopped-flow techniques, effects on the microscopic rates were also studied. A number of the most interesting variants were also crystallized and their structures were determined by X-ray crystallography (A2, pdb ID: 5o8q, A2C2 pdb ID: 5o8h, A2C3 pdb ID: 5o9f, and A2C2B1 pdb ID: 5o9d) (Figure 22C–F).

Kinetic Reaction Mechanism

In wild type ADH-A, the rate determining steps for alcohol oxidation or corresponding ketone reduction have been proposed to be the rates of a conformational changes in the enzyme-nucleotide binary complex, such as E*-NAD(H) → E-NAD(H), similarly to the situation described for the horse-liver enzyme. This was inferred from observations of the NADH concentration dependency of the transient rates of nucleotide binding to the active site in wild type ADH-A. Here, we reiterated the same studies of nucleotide binding and instead concluded that ADH-A catalyzed oxidation or reduction proceeds via a simple one-to-one binding and dissociation model of an ordered five-step bi-bi mechanism (Figures 22 and 23). The off-rates of NADH (k₅) and NAD⁺ (k⁻₁) are in agreement with the turnover numbers for (S)-1-phenylethanol and acetophenone (Table 10 and 12), and couples the rate limitation to coenzyme release, in the reactions with these preferred substrates.

Oxidation of 2-propanol is catalyzed by wild type ADH-A with similar turnover number to that with (S)-1-phenylethanol. The deuterium kinetic isotopic
effect (KIE) determined with 2-propanol-d\textsubscript{1} differed on \(k_{\text{cat}}\) is low (1.5), which further supports the notion that step(s) downstream of the isotope sensitive oxidation step is primarily rate limiting for turnover. In addition, the KIE on the rate of oxidation, \(k_3\), is substantially larger (3.6) (Table 11), fitting the same argument (Table 12). Similarly, the larger KIE for \(k_{\text{cat}}/K_M\) also suggested that a rate-limiting step is downstream of the ketone release.

Figure 22. Observed rates as a function of NAD\textsuperscript{+}/NADH concentration by wild type and different ADH-A variants. A: filled symbols, wild type with NAD\textsuperscript{+} (circles), or NADH (squares). Unfilled symbols, are binding data of NADH from different variants; A2 (circles), A2C2 (squares), A2C3 (triangles), A2C2B1 (diamonds) and A2C2B1 (inverted triangles). The extracted values of binding and dissociation rates is shown in Table 10. B: Example fluorescence tracks recorded during the initial stage of coenzyme (NADH) binding with wild type and A2C2B2 variant. These tracks were recorded after mixing 10 \(\mu\)M NADH binding with 1.0 \(\mu\)M respective ADH-A variants in the stopped-flow apparatus. The pre-equilibrium burst of binding and release was followed by linear equilibrium phase.

Figure 23. Current model of the kinetic mechanism of alcohol oxidation catalyzed by ADH-A. Reaction proceeds with NAD\textsuperscript{+} -binding (\(k_1\)) to the free enzyme, which lead to a binary enzyme-NAD\textsuperscript{+} -complex (E-NAD\textsuperscript{+}). To this complex, the alcohol substrate (R-OH; \(k_2\)) binds. Thereafter, oxidation occurs (\(k_3\)), and the product (ketone) is released (\(k_4\)), followed by final release of NADH (\(k_5\)) from the enzyme. The rates of all the steps subsequent to formation of ternary complex (\(k_3\) to \(k_5\)) will contribute to \(k_{\text{cat}}\), whereas \(k_{\text{cat}}/K_M\) includes \(k_2\) to \(k_4\).
Table 10. Microscopic kinetic rates and equilibrium dissociation constants for nucleotide binding and release. Dissociation constant for NAD$^+$ ($K_1$) is calculated by $k_1/k_1$ and the dissociation constant for NADH ($K_5$) is calculated by $k_5/k_5$. n.d., not detected fluorescence, due to loss of fluorophores from Y294F and W295A mutations

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$K_1$ (µM)</th>
<th>$k_1$ (s$^{-1}$µM$^{-1}$)</th>
<th>$k_1$ (s$^{-1}$)</th>
<th>$k_5$ (s$^{-1}$µM$^{-1}$)</th>
<th>$k_5$ (s$^{-1}$)</th>
<th>$K_5$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>NAD$^+$</td>
<td>12±0.8</td>
<td>2.7±0.06</td>
<td>31±2</td>
<td>51±6</td>
<td>5.1±0.1</td>
<td>10±1</td>
</tr>
<tr>
<td>A1</td>
<td>NAD$^+$</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>270±20</td>
<td>1.7±1</td>
<td>16±2</td>
</tr>
<tr>
<td>A2C2</td>
<td>NAD$^+$</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>260±4</td>
<td>9.5±0.2</td>
<td>27±0.6</td>
</tr>
<tr>
<td>A2C3</td>
<td>NAD$^+$</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>200±60</td>
<td>27±4</td>
<td>7.4±2</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>NAD$^+$</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>65±6</td>
<td>26±0.2</td>
<td>2.5±0.2</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>NAD$^+$</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>110±6</td>
<td>24±0.2</td>
<td>4.4±0.2</td>
</tr>
</tbody>
</table>

The lower KIE on $k_{cat}$ indicates that $Dk_{cat}/K_{0.5}$ is expressed in reaction step(s) from the E-NAD$^+$ complex to the release of the first product acetone (Table 11). Therefore, in an ordered mechanism, this supports that the later step of NADH release with rate constant $k_5$, is the main rate limiting step for the overall turnover.

Table 11. Deuterium kinetic isotopic effect on catalyzed 2-propanol oxidation with wild type and selected ADH-A variants. b, Hill-coefficient, if $n = 1$, then the $K_{0.5}$ becomes $K_M$. n.s., refers to no enzyme saturation with in the usable substrate saturation range.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$k_3$ (s$^{-1}$)</th>
<th>$Dk_3$</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$Dk_{cat}$</th>
<th>$k_{cat}/K_{0.3}b$ (s$^{-1}$mM$^{-1}$)</th>
<th>$Dk_{cat}/K_{0.5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>2-propanol</td>
<td>430±90</td>
<td>3.6±0.8</td>
<td>85±1</td>
<td>1.5±0.05</td>
<td>3.0±0.1</td>
<td>2.8±0.03</td>
</tr>
<tr>
<td></td>
<td>2-propanol-2-d</td>
<td>120±10</td>
<td>85±1</td>
<td>1.4±0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>25±1</td>
<td>3.1±0.1</td>
<td>0.045±0.003</td>
<td></td>
<td>2.9±0.3</td>
</tr>
<tr>
<td>A2</td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>8.2±0.2</td>
<td>0.016±0.0007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2C2</td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>40±4</td>
<td>3.1±0.05</td>
<td>0.024±0.004</td>
<td></td>
<td>1.6±0.3</td>
</tr>
<tr>
<td></td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>40±4</td>
<td>3.1±0.05</td>
<td>0.024±0.004</td>
<td></td>
<td>1.6±0.3</td>
</tr>
<tr>
<td>A2C2</td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>3±0.7</td>
<td>0.015±0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-propanol-2-d</td>
<td>-</td>
<td>n.s.</td>
<td>0.0033</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First Generation

Variants A1 (W295A) and A2 (Y294F, W295A – shown in Figure 21), were purified and characterized. The results showed that the W295A mutation was responsible for most of the observed effects on enantioselectivity. The substitution introduced an additional pocket in the active site cavity due to the absence of tryptophan indole ring. This change facilitated more productive binding of (R)-1-phenylethanol, observed as an increase in $k_{cat}$ (4.9-fold) in these variants (Table 12). However, the increase in the turnover number is also accompanied by 10- and 8-fold increases in $K_M$ for A1 and A2, respectively. Therefore, the resulting values of $k_{cat}/K_M$ were essentially unchanged. The increase in $K_M$ and $k_{cat}$ can be explained by a decrease in nonproductive binding.
and providing a possibility for the substrate to probe a larger space, for productive binding, facilitating the increased turnover number. Furthermore, in the A2 variant NADH release is five-fold faster as compared to the wild type \( k_5^{A2}/k_5^{wt} \) in Table 10) and was shown to be no longer rate limiting for catalytic turnover. The release of NADH further concluded by deuterium kinetic isotopic effect measurements (Table 11), showing that \( k_{cat} \) (3.1) and \( k_{cat}/K_M \) (2.9). These values are very close to the values the KIE of the oxidation step \( k_3 \), determined in the wild type ADH-A catalyzed reactions. If assuming that the intrinsic isotopic effect is unchanged by the mutations, these results suggests that the oxidation step is now main rate-limiting step in the A2 variant.

The A2 variant have a shifted enantiopreference to favor oxidation of \((R)-1\)-phenylethanol over the corresponding \((S)-1\)-phenylethanol by 2.8-fold. However, the same enantiomeric ratio of the alcohol products was not observed following A2 catalyzed reduction of acetophenone. This reaction resulted in a racemic mixture of \((R)-1\)-, and \((S)-1\)-phenylethanol. This may be due to a lower stringency in the substrate binding geometry for the planar acetophenone in the hydride transfer from NADH, which in this mutant can occur to either face of the carbonyl carbon. In contrast, to achieve efficient oxidation of the alcohol, which must be precisely positioned to allow the hydride to be transferred to the C-4 atom of nicotinamide ring. This proposal was supported by docking experiments performed on the wild type enzyme which revealed that the position of acetophenone cannot be bound in a manner that would allow for hydride addition in \( si \)-face, due to clashes between the phenyl substituent and the indole group of tryptophan (W295). This limitation is removed in the A2 variant due to the W295A substitution.

The A2 variant preference towards \((R)-1\)-phenylethanol is primarily caused by a 1300-fold decrease in the activity with \((S)-1\)-phenylethanol (Table 12), presumably due to an increase in nonproductive binding of this enantiomer. This was also supported by docking simulations, which suggest that the positioning of the substrate is restricted by narrow active site cavity and coordination of the hydroxy oxygen to the active site zinc. Therefore, in the wild type only \((S)-1\)-phenylethanol can be bound and oriented in a manner in which hydride transfer can takes place from C-1 of the alcohol to C-4 of the nicotinamide ring, whereas \((R)-1\)-phenylethanol is bound in nonproductive orientation due to the ‘wrong’ positioning of the C-1 hydrogen. Due to the larger active site cavity, the \((R)-\)enantiomer could bind productively in the A2 variant, whereas \((S)-1\)-phenylethanol was bound nonproductively, in the most preferred docking pose. The second position in subsite A, Y294 was either mutated to phenylalanine in A2 or unchanged as tyrosine in A1. These aromatic characteristics probably facilitated the \( \pi-\pi \) stacking interaction between the aromatic side chains of Y294 (or F294) and F286 from the neighboring subunit contributing to stabilization of the dimer interface.
Figure 24. Crystal structures of ADH-A wild type and evolved variants. A: superimposed cartoon representation of the subunit structures of wild type and variants. Active sites amino acid residues in B: wild type (brown) and evolved variants C: A2 (violet purple), D: A2C2 (sky-blue), E: A2C3 (yellow) and F: A2C2B1 (salmon) from three successive generations, respectively.
Table 12. Steady-state kinetic parameters of ADH-A wild type and evolved variants selected for improved (R)-phenylethanol. If the Hill-coefficient n, is >1, then the $K_M$ becomes $K_{0.5}$ which is equal to the $[S]$ that results in half maximum reaction velocity, and also $k_{cat}/K_{0.5}$. *, adapted data from reference [119]. n.s., refers to no enzyme saturation with in the usable substrate saturation range.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$(R)$-1-phenylethanol $k_{cat}$ (s$^{-1}$)</th>
<th>$(R)$-1-phenylethanol $K_M$ (mM)</th>
<th>$(R)$-1-phenylethanol $k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
<th>(S)-1-phenylethanol $k_{cat}$ (s$^{-1}$)</th>
<th>(S)-1-phenylethanol $K_M$ (mM)</th>
<th>(S)-1-phenylethanol $k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
<th>Acetophenone $k_{cat}$ (s$^{-1}$)</th>
<th>Acetophenone $K_M$ (mM)</th>
<th>Acetophenone $k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
<th>2-propanol $k_{cat}$ (s$^{-1}$)</th>
<th>2-propanol $K_M$ (mM)</th>
<th>2-propanol $k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>0.45 ±0.005*</td>
<td>0.94 ±0.04*</td>
<td>0.48 ±20*</td>
<td>0.63 ±30*</td>
<td>0.0037</td>
<td>36</td>
<td>1.2 ±0.8*</td>
<td>30</td>
<td>85</td>
<td>22</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>2.2 ±0.1</td>
<td>9.7 ±0.7</td>
<td>0.23 ±0.2</td>
<td>0.48 ±0.2</td>
<td>2.7 ±0.1</td>
<td>0.18 ±0.2</td>
<td>1.3 ±0.2</td>
<td>2.8 ±0.2</td>
<td>0.18 ±0.2</td>
<td>9.0 ±0.2</td>
<td>0.020 n.s</td>
<td>n.s.</td>
</tr>
<tr>
<td>A2</td>
<td>2.2 ±0.04</td>
<td>7.7 ±0.3</td>
<td>0.28 ±0.01</td>
<td>0.26 ±0.01</td>
<td>2.5 ±0.1</td>
<td>0.10 ±0.1</td>
<td>2.8 ±0.2</td>
<td>0.18 ±0.2</td>
<td>9.0 ±0.2</td>
<td>0.020 n.s</td>
<td>550 ±0.045</td>
<td></td>
</tr>
<tr>
<td>A2C1</td>
<td>2.3 ±0.1</td>
<td>13 ±0.1</td>
<td>0.18 ±0.02</td>
<td>0.12 ±0.02</td>
<td>3.0 ±0.1</td>
<td>0.040 ±0.02</td>
<td>4.5 ±0.1</td>
<td>0.006 ±0.006</td>
<td>0.9 ±0.1</td>
<td>0.001 ±0.001</td>
<td>25 ±0.01</td>
<td>30 ±0.003</td>
</tr>
<tr>
<td>A2C2</td>
<td>4.0 ±0.2</td>
<td>10 ±0.9</td>
<td>0.40 ±0.04</td>
<td>0.22 ±0.06</td>
<td>2.3 ±0.3</td>
<td>0.096 ±0.01</td>
<td>4.2 ±0.1</td>
<td>0.001 ±0.001</td>
<td>0.9 ±0.1</td>
<td>0.004 ±0.004</td>
<td>40 ±0.3</td>
<td>1600 ±0.024</td>
</tr>
<tr>
<td>A2C3</td>
<td>20 ±7</td>
<td>100 ±0.2</td>
<td>0.20 ±0.02</td>
<td>0.67 ±0.04</td>
<td>3.1 ±0.3</td>
<td>0.22 ±0.1</td>
<td>0.90 ±0.01</td>
<td>0.73 ±0.03</td>
<td>8.3 ±0.2</td>
<td>0.089 n.s</td>
<td>n.s.</td>
<td>0.040</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>3.4 ±0.2</td>
<td>3.6 ±0.3</td>
<td>0.45 ±0.05</td>
<td>0.26 ±0.06</td>
<td>4.2 ±0.1</td>
<td>0.061 ±0.01</td>
<td>7.4 ±0.1</td>
<td>0.001 ±0.001</td>
<td>0.3 ±0.1</td>
<td>0.009 ±0.008</td>
<td>4.4 ±0.1</td>
<td>n.s.</td>
</tr>
<tr>
<td>A2C2B2</td>
<td>9.3 ±2</td>
<td>80 ±0.8</td>
<td>0.12 ±0.02</td>
<td>0.55 ±0.05</td>
<td>16 ±0.1</td>
<td>0.034 ±0.01</td>
<td>3.5 ±0.1</td>
<td>0.25 ±0.01</td>
<td>4.6 ±0.05</td>
<td>0.005 ±0.008</td>
<td>4.4 ±0.1</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
Second Generation
The isolated A2C offspring, resulting from saturation mutagenesis at the “C” subsite that were scored as potential hits, all retained I271. This may be due to the residue being in VdW contact with the nicotinamide ring of the coenzyme and may fulfill a role in maintaining basic catalytic function. Isolated variants from A2C1–C3 (Table 9) were further analyzed for the structure-function properties. Both variants A2C1 and A2C2 contained an F43H substitution, whereas A2C2 also contained an additional mutation H39Y. This substitution was not originally encoded into the initial gene libraries. Thus, the H39Y mutation was viewed as a PCR artifact (Figure 24 A and B). The H39Y substitution results in the loss a hydrogen bond between H39 side-chain and a phosphate oxygen of the coenzyme. Residue H39 residue is located in a region of higher crystallographic B-factors, indicating a larger degree of mobility. Still the loss of a hydrogen bond could potentially influence NADH binding and release. The difference in the binding behavior of the parent A2 and A2C2, however, is negligible (Table 10).

The mutation F43H in A2C1 has a relatively small effects on substrate preference, causing a three-fold decrease in activity with (S)-1-phenylethanol, and improved selectivity for (R)-1-phenylethanol. The crystal structure of A2C2 revealed a possible role of this substitution; H43 participates in a new hydrogen bond interaction with the 2'-hydroxyl of the coenzyme ribose. Although, this new interaction can be expected to affect coenzyme binding and release, the interaction kinetics is similar to that of parent A2. Interestingly, H43 in A2C2 is located in the corresponding position to H51 in the horse-liver ADH, a residue that has been assigned to take part in the proton relay chain to channel the alcohol proton during catalyzed oxidation in that enzyme.

In A2C3 variant, instead an F43S substitution was present in addition to the H39Y mutation. This single residue difference in comparison with A2C2, resulted in five-fold higher turnover with (R)-1-phenylethanol. The increase in turnover number was also accompanied by a concomitant increase in $K_M$, resulting in an unaffected overall efficiency (Table 12). The crystal structure of A2C3 (Figure 24E) illustrated that the replacement of S43 cannot directly form a hydrogen bond with the coenzyme, as in the case of H43. However, a corresponding interaction can still be made via a bridging water molecule. Moreover, A2C3 the higher turnover number with (R)-1-phenylethanol indicates that enzyme-coenzyme interactions between non-catalytic residues can contribute to the catalytic efficiency.

Third Generation:
the variant A2C2 was chosen to parent a third round of saturation mutagenesis in at the “B” site (Figure 21). The variants scored as hits all retained L119 but...
contained different substitutions at position Y54, Y54F in A2C2B1 and Y54W in A2C2B2 (Table 9). These two variants A2C2B1 and A2C2B2 were purified and characterized. The kinetic measurements displayed similar behavior in the case of variant A2C2B1, as compared to its parent A2C2 (Table 12), but with an increased $E$-value for (R)-1-phenylethanol. The crystal structure of A2C2B1 showed that F54 orients towards the active site (Figure 24F).

B-Factor Analysis
To identify the regions of flexibility regions we compared the crystallographic temperature factors (B-Factors) of ADH-A and the crystallized variants. Four distinctive regions displaying relatively higher B-factor values could be identified involving residues from 6–13, 37–57, 108–120 and 331-339. Although these regions are spread out over the entire amino acid sequence, they come together in the folded structure and contribute to a major part of enzyme active site. Therefore, three of these regions of apparent higher intrinsic mobility were harboring the mutated residues in the different libraries. Hence, a contribution to the obtained differences in kinetic parameters and binding rates, can possibly be attributed to alterations in structural dynamic properties.

Conclusions
The main objective of this work was to generate the ADH-A variants with preference for (R)-1-phenylethanol over (S)-1-phenylethanol. Preference for (R)-enantiomer was demonstrated in hits from all the three generations, where the best variant was isolated from third generation of A2C2B1 displayed 7.4-fold preference for the (R)-enantiomer. In comparison with the wild type enzyme this variant displayed a 2000-fold higher $E$-value for (R)-1-phenylethanol. The significant change was attributed to structural changes via W295A substitution that created a new cavity in the active site, which facilitated productive binding of (R)-1-phenylethanol to a larger degree.

It was also observed that an increase in $k_{cat}$ was paralleled by an increase in $K_M$, thus resulting in very small changes in overall catalytic efficiency. However, from an industrial point of view, the increase in $k_{cat}$ is more valuable than higher $k_{cat}/K_M$ values, since in an industrial set-up higher reactant (substrate) concentrations are expected to be applied leading to enzyme saturation. Furthermore, the higher $K_M$ values can also be beneficial, since it reduces the risk of substrate inhibition. The parallel increases in $k_{cat}$ and $K_M$ can be attributed to a decrease in nonproductive substrate binding. A combination of lower degree of nonproductive binding and increased release rates of NADH can enhance the increase the turnover number.
Paper II

Wild type ADH-A exhibits a 270-fold preference towards the (S)-enantiomer of 1-phenylethanol over the corresponding (R)-enantiomer. When challenged with chiral vicinal 1,2-diols as putative substrates wild type ADH-A displayed preference for (R)-1-phenyl-1,2-ethanediol over (S)-1-phenyl-1,2-ethanediol. This apparent shift in enantioselectivity going from 1-phenylethanol to 1-phenyl-1,2-ethanediol is simply an effect of the change in the order of priorities of the substitutions bonded to the stereogenic-center. However, the wild type ADH-A catalyzed oxidation of vicinal 1,2-diols is relatively inefficient, (R)-1-phenyl-1,2-ethanediol is oxidized into 2-hydroxyacetophenone approximately 2600-fold slower as compared to the reaction with (S)-1-phenylethanol, forming acetophenone, if comparing the values of $k_{cat}/K_M$. Although the catalytic activity is comparably modest, the formed $\alpha$-hydroxy ketone (acyloin) product underlines the stringent regioselectivity of this enzyme, only catalyzing oxidation of the sec-alcohol group.\textsuperscript{112,119}

The oxidation of vicinal 1,2-diols involves the removal of an asymmetric center into a prochiral ketone. However, the corresponding acyloin is an attractive building block for the production of important carbon–carbon bonded derivatives, such as chiral auxiliaries, natural products, fine chemicals and pharmaceuticals. Therefore, in this study we tried to improve the catalytic activity of ADH-A with vicinal 1,2-diol, such as (R)-1-phenyl-1,2-ethanediol (Figure 25), by generating variants by the same approach as described in Paper I. ADH-A variants of improved catalytic activities may be directly employed in multi-step enzyme catalyzed reactions (Figure 26), thus, a starting reactant can be cheaper racemic styrene oxide, which can be transformed into a more valuable acyloin.

Figure 25. Alcohols and ketones: (S)-1-phenylethanol, (R)-1-phenyl-1,2-ethanediol, acetophenone and 2-hydroxyacetophenone “1–4” were tested as substrates in this study.

Figure 26. (R)-2, was produced by epoxide hydroxylase (StEH1) catalyzed hydrolysis of corresponding styrene oxide. Desired, oxidation of (R)-2 with laboratory evolved ADH-A variants selectively oxidize the “C-1” sec-alcohol to the corresponding 2-hydroxyacetophenone, 4, respectively.
Library Generation and Evolutionary Routes

To optimize the catalytic efficiency of ADH-A with (R)-1-phenyl-1,2-ethanediol, we re-mined the enzyme libraries described in Paper I. Searching through the same three libraries (A–C), for ADH-A variant of apparently higher activity with (R)-1-phenyl-1,2-ethanediol (Figure 21 and 27B).120

![Figure 27. Crystal structures of ADH-A wild type and evolved variants. A: superimposed cartoon representation of the subunit structures of wild type and variants. Active sites amino acid residues in B: wild type (brown) and evolved variants C: C1 (light-teal), and D: C1B1 (sand) from two successive generations, respectively.]

In this case, the A or B libraries did not provide any candidate hits, but during the searching through the C library we observed several apparently improved variants. The most active variant C1 (F43H) was subsequently isolated and further characterized. (This mutation was also found in the second-generation variants A2C1 and A2C2, as described in Paper I.) The C1 variant was chosen to parent a second round of saturation mutagenesis, re-visiting both the A and B site. No hits were identified from the C1A library, but in the C1B library
we observed several improved variants. The most active variant C1B1 (F43H, Y54L) was subsequently isolated and further characterized. In addition, to the kinetic analysis, the tertiary structures of C1 and C1B1 were determined by X-ray crystallography (C1, pdb ID: 6ffx, and C1B1 pdb ID: 6ffz) (Figure 27C, and D).

Kinetic Effects on Variants
As described in Paper I, the rate limiting step for the catalyzed oxidation of (S)-1-phenylethanol is the release of NADH ($k_5$, in Table 14 and Figure 23). Wild type ADH-A displays an NADH release rate that is 70-fold faster than $k_{cat}$ for (R)-1-phenyl-1,2-ethanediol. Assuming that oxidation of (R)-1-phenyl-1,2-ethanediol still follows an ordered bi-bi mechanism, the reason for the lower $k_{cat}$ can be due to nonproductive substrate binding in the active site of the enzyme. The isolated variants displayed increases in $k_{cat}$ for (R)-1-phenyl-1,2-ethanediol, 2.6-fold for C1 (F43H), and 7.6-fold for C1B1 (F43H, Y54L), respectively, as compared to the wild type (Table 13). Similar to the previous findings in Paper I, the increase in $k_{cat}$ is accompanied by parallel increases in $K_M$, thus the overall catalytic efficiencies $k_{cat}/K_M$ are unchanged in C1 and C1B1. This, as in the previous case, indicates that an increase in productive binding of (R)-1-phenyl-1,2-ethanediol in the ternary complex may be the underlying reason for the improved turnover number in the variants.

<table>
<thead>
<tr>
<th>Enzyme/Substrates</th>
<th>$k_{cat}$ ($s^{-1}$)</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}/K_M$ ($s^{-1}\times mM$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type / (S)-1</td>
<td>80±20</td>
<td>0.63±0.05</td>
<td>130±30</td>
</tr>
<tr>
<td>wild type / (R)-2</td>
<td>0.73±0.01</td>
<td>17±0.6</td>
<td>0.044±0.0008</td>
</tr>
<tr>
<td>wild type / 3</td>
<td>36±0.8</td>
<td>1.2±0.09</td>
<td>30±2</td>
</tr>
<tr>
<td>wild type / 4</td>
<td>2.04±0.08</td>
<td>3.7±0.4</td>
<td>0.55±0.03</td>
</tr>
<tr>
<td>C1 / (R)-2</td>
<td>1.94±0.05</td>
<td>37±2</td>
<td>0.05±0.001</td>
</tr>
<tr>
<td>C1 / 4</td>
<td>19±3</td>
<td>36±6</td>
<td>0.53±0.02</td>
</tr>
<tr>
<td>C1B1 / (R)-2</td>
<td>5.5±0.4</td>
<td>120±20</td>
<td>0.03±0.003</td>
</tr>
<tr>
<td>C1B1 / 4</td>
<td>8.8±1</td>
<td>20±4</td>
<td>0.44±0.04</td>
</tr>
</tbody>
</table>

Reduction of 2-hydroxyacetophenone catalyzed by the C1 variant displayed a 10-fold increase in $k_{cat}$, and $K_M$, whereas C1B1 displayed a 4.3-fold increase as compared to the wild type (Table 13). Again, the overall values of $k_{cat}/K_M$ are approximately the same for both the variants and wild type. In addition, the rates of alcohol oxidation ($k_3$) with (R)-1-phenyl-1,2-ethanediol to 2-hydroxyacetophenone are significantly lower, as compared to the oxidation of (S)-1-phenylethanol, although still an order of magnitude faster than the values of $k_{cat}$ (Table 14). The positive effect of $k_{cat}$ (R)-1-phenyl-1,2-ethanediol must be an effect of the combination of the H43 and L54 substitutions in C1B1. This suggests a synergistic effect between the H43 and L54...
substitutions, since without the first insertion of the F43H mutation, the added positive effect of the Y54L substitution could apparently not penetrate.

Table 14. Kinetic parameters of ADH-A wild type and evolved variants (C1 and C1B1) microscopic rates of the alcohol oxidation step \((k_3)\) and NADH release \((k_5)\). Data adapted from references 119 and 120.

<table>
<thead>
<tr>
<th>Enzyme/Substrates</th>
<th>(k_3 (s^{-1}))</th>
<th>(k_5 (s^{-1}))</th>
<th>(k_{cat} (s^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type / (S)-1</td>
<td>630±40</td>
<td>51±6</td>
<td>80±20</td>
</tr>
<tr>
<td>wild type / (R)-2</td>
<td>42±8</td>
<td>51±6</td>
<td>0.73±0.01</td>
</tr>
<tr>
<td>C1 / (S)-1</td>
<td>-</td>
<td>76±20</td>
<td>110±7</td>
</tr>
<tr>
<td>C1 / (R)-2</td>
<td>25±2</td>
<td>76±20</td>
<td>1.9±0.05</td>
</tr>
<tr>
<td>C1B1 / (S)-1</td>
<td>340±20</td>
<td>93±10</td>
<td>120±6</td>
</tr>
<tr>
<td>C1B1 / (R)-2</td>
<td>58±10</td>
<td>93±10</td>
<td>5.5±0.4</td>
</tr>
</tbody>
</table>

The single substitution of F43H in C1, resulted in a structural alteration that was, as mentioned, also observed in the second-generation variant A2C1 and A2C2 in Paper I. Superimposition of C1 and A2C2 structures revealed that H43 residue in both variants were positioned identically. The imidazole side chain of H43 participates in a new hydrogen bond interaction with the 2'-hydroxyl of the cofactor ribose. In this conformation, the imidazole ring of H43 residue flips away from the substrate binding domain, causing an increase in the volume of the active site cavity as compared to F43 in the wild type. Clearly, the single F43H mutation is responsible for the increase in \(k_{cat}\) for (R)-1-phenyl-1,2-ethanediol by the C1 variant. However, the significance of this structural changes regarding the new interactions between enzyme and cofactor remains unclear. In C1B1, the F43H substitution makes the same interaction with the cofactor and together with the Y54L substitution located approximately perpendicular to H43, caused a further enlargement of the active site cavity, resulting in slightly further increased \(k_{cat}\) in the oxidation and reduction reaction.

Conclusion
The main objective of this work\textsuperscript{122} was to generate ADH-A variants with increased catalytic activity with the vicinal diol analog of (S)-1-phenylethanol, (R)-1-phenyl-1,2-ethanediol, to enable efficient synthesis of the corresponding \(\alpha\)-hydroxy acyloin, and vice versa, for asymmetric production of the chiral alcohol. The resulting C1B1 variant can be employed in multi-step enzymatic reaction pathway. Both the C1 (F43H) and C1B1 (F43H, Y54L) displayed increased turnover number with (R)-1-phenyl-1,2-ethanediol, accompanied by parallel increases in \(K_M\) resulting in unchanged overall catalytic efficiencies \((k_{cat}/K_M)\). However, as also discussed in the conclusions to Paper I, considering a putative industrial application, a higher \(k_{cat}\) assumed to of higher practical value over low \(K_M\) values (or higher \(k_{cat}/K_M\) values), since higher substrate concentrations are most probably applied leading to enzyme saturation.
Paper III

The strong regioselectivity in the oxidation of 1-phenyl-1,2-ethanediol triggered us to further probe the regioselectivity. In this work we challenged the same enzyme with a substrate containing two sec-alcohol (C-1 and C-2) functions such as 1-phenylpropane-(1R,2S)-diol ((1R,2S)-3) and the cognate diketone (8) (Figure 28). Not surprisingly, considering the structural similarities between (1R,2S)-3 and (R)-2, the oxidation of (1R,2S)-3 is relatively poor as compared to (S)-1-phenylethanol ((S)-1)). As described in Paper II, ADH-A catalyzed oxidation of vicinal 1,2-diols results in the removal of an asymmetric center into a prochiral ketone. Therefore, in this study we focused on generating variants that displayed improved regioselectivity in oxidation with a (1R,2S)-3, by the same approach as described in Paper I. ADH-A variants of improved catalytic activities may be directly employed in multi-step enzyme catalyzed reactions (Figure 30), thus, a starting reactant can be cheaper racemic styrene oxide, which can be transformed into a more valuable acyloins, e.g. to produce chiral amino alcohols.

![Diagram of alcohols and ketones](image)

Figure 28. Alcohols and ketones: (S)-1-phenylethanol, (R)-1-phenylethanol, (R)-1-phenyl-1,2-ethanediol, (S)-1-phenyl-1,2-ethanediol, 1-phenylpropane-(1R,2S)-diol, acetophenone, 2-hydroxyacetophenone “1–5” and 1-phenylpropane-1,2-dione “8” were tested as substrates in this study. Acyloins, 1-phenyl-(2S)-hydroxy-2-propanone, 1-phenyl-(1R)-hydroxy-2-propanone “6 and 7” are the oxidation products from (1R,2S)-3.

When challenged with (1R,2S)-3 as substrate, (wild type ADH-A preferentially oxidizes the benzylic C-1 carbon, producing (S)-6 in a ratio of 3.6:1 over (R)-7, as determined by NMR (Table 16, Figure 31). The corresponding reduction of 8 displayed high $k_{cat}$ ($62 \text{ s}^{-1}$) and the $k_{cat}/K_{0.5}$ values, indicated further emphasized this diketone as a preferred substrate. Product analysis showed that reduction of 8 is preferentially (20:1) at C-2 forming acyloin 6 (Table 16).

Library Generation and Evolutionary Routes

To optimize the catalytic efficiency of ADH-A with (1R,2S)-3, we re-mined the enzyme libraries described in Paper I. Searching through the same A–B
libraries, for ADH-A variant of apparently higher activity with \((1R,2S)\)-3 (Figure 21 and 29B).\(^{120}\)

Figure 29. Crystal structures of ADH-A wild type and evolved variants. A: superimposed cartoon representation of the subunit structures of wild type and variants. Active sites amino acid residues in B: wild type (brown) and evolved variants C: B1 (light blue), and D: B1F4 (raspberry) from three successive generations respectively.

In this case, the A library did not provide any candidate hits, but during the search through the B library we observed one candidate variant that displayed apparently higher activity for \((1R,2S)\)-3, as compared to the parent wild type. The most active variant B1 (Y54G and L119Y) was isolated and further characterized. Thereafter, the B1 variant was chosen to parent a second round of saturation mutagenesis with slight modifications, generating a B1F library. The “F” site, contained two residues, F43, previously also targeted in the “C” site and F282\(^{\#}\) from the neighboring subunit of the dimer. In the B1F library
we observed several improved variants. The most active variant B1F4 (F43T, Y54G, L119Y and F282W*) was subsequently isolated and further characterized. In addition, to the kinetic analysis, the tertiary structures of B1 and B1F4 were determined by X-ray crystallography (B1, pdb ID: 5od3, and B1F4 pdb ID: 6fg0) (Figure 29C and D).

**Kinetic Effects on Variants**

Wild type ADH-A catalyzed oxidation of (1R,2S)-3 is relative 170-fold lower, if comparing $k_{cat}$, then the monohydroxylated (S)-1-phenylethanol, and also displayed negative cooperativity with a Hill-coefficient 0.75. This is a typical suggestion the transition to a less active enzyme form with increasing concentration of substrate. Variants, B1 (Y54G and L119Y) and B1F4 (F43T, Y54, L119Y and F282W*) displayed 2-fold and 4.8-fold increase in $k_{cat}$ for (1R,2S)-3 as compared to the parent wild type. Both these enzymes shifted from negative to positive cooperativity for this substrate with Hill-coefficients of 1.2 and 1.4, respectively (Table 15).

**Table 15.** Steady-state kinetic parameters of ADH-A wild type and evolved variants selected for improved (1R,2S)-3 and corresponding cognate di-ketone 8. $n^*$, Hill-efficient, if $n = 1$, then the $K_{0.5}$ becomes $K_M$ which is equal to the [S] that results in half maximum reaction velocity, and also $k_{cat}/K_{0.5}$. Data adapted from reference 118. *b*, amino acid residues from neighboring dimer subunit.

<table>
<thead>
<tr>
<th>Substrates</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_{0.5}$ (mM)</th>
<th>$k_{cat}/K_{0.5}$ (s$^{-1}$mM)</th>
<th>$n^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(S)-1</strong></td>
<td>80±20</td>
<td>0.63±0.05</td>
<td>130 000±30 000</td>
<td>(1)</td>
</tr>
<tr>
<td><strong>(R)-2</strong></td>
<td>0.73±0.01</td>
<td>17±0.6</td>
<td>44±8</td>
<td></td>
</tr>
<tr>
<td><strong>(S)-2</strong></td>
<td>0.0094±0.0005</td>
<td>3.0±0.7</td>
<td>3.1±0.6</td>
<td></td>
</tr>
<tr>
<td>**(1R,2S)-3</td>
<td>0.48±0.01</td>
<td>11±1</td>
<td>43±5</td>
<td>0.75±0.03</td>
</tr>
<tr>
<td>4</td>
<td>36±0.8</td>
<td>1.2±0.09</td>
<td>30 000±2 000</td>
<td>(1)</td>
</tr>
<tr>
<td>5</td>
<td>2.0±0.08</td>
<td>3.7±0.4</td>
<td>550±30</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>62±0.6</td>
<td>0.31±0.01</td>
<td>200 000±9 000</td>
<td>1.3±0.06</td>
</tr>
<tr>
<td><strong>B1 (Y54G, L119Y)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(S)-1</strong></td>
<td>11±1</td>
<td>32±6</td>
<td>340±30</td>
<td>(1)</td>
</tr>
<tr>
<td><strong>(R)-2</strong></td>
<td>0.17±0.008</td>
<td>130±10</td>
<td>1.3±0.06</td>
<td>(1)</td>
</tr>
<tr>
<td><strong>(S)-2</strong></td>
<td>0.039±0.001</td>
<td>17±2</td>
<td>2.3±0.2</td>
<td>(1)</td>
</tr>
<tr>
<td>**(1R,2S)-3</td>
<td>0.98±0.03</td>
<td>33±2</td>
<td>30±2</td>
<td>1.2±0.06</td>
</tr>
<tr>
<td>4</td>
<td>7.3±0.6</td>
<td>4.8±1</td>
<td>1 500±300</td>
<td>(1)</td>
</tr>
<tr>
<td>5</td>
<td>0.54±0.02</td>
<td>15±0.9</td>
<td>36±1</td>
<td>(1)</td>
</tr>
<tr>
<td>8</td>
<td>8.5±0.5</td>
<td>0.32±0.009</td>
<td>260 000±8 000</td>
<td>0.93±0.02</td>
</tr>
<tr>
<td><em><em>B1F4 (F43T, Y54G, L119Y, F282W</em>)</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(S)-1</strong></td>
<td>37±3</td>
<td>73±10</td>
<td>500±100</td>
<td>0.77±0.02</td>
</tr>
<tr>
<td><strong>(R)-2</strong></td>
<td>0.45±0.006</td>
<td>140±4</td>
<td>3.2±0.004</td>
<td>(1)</td>
</tr>
<tr>
<td><strong>(S)-2</strong></td>
<td>0.072±0.008</td>
<td>60±20</td>
<td>1.2±0.04</td>
<td>0.79±0.06</td>
</tr>
<tr>
<td>**(1R,2S)-3</td>
<td>2.3±0.04</td>
<td>35±1</td>
<td>66±3</td>
<td>1.4±0.05</td>
</tr>
<tr>
<td>4</td>
<td>26±0.6</td>
<td>8.5±0.5</td>
<td>3 000±200</td>
<td>1.2±0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.98±0.2</td>
<td>25±1</td>
<td>38±1</td>
<td>(1)</td>
</tr>
<tr>
<td>8</td>
<td>160±2</td>
<td>0.51±0.02</td>
<td>320 000±2 000</td>
<td>(1)</td>
</tr>
</tbody>
</table>
On the other hand, the oxidation rate ($k_3$) for $(1R,2S)$-3 has also increased approximately 4-folds, and exceeds the value of $k_{cat}$ value by an order of magnitude similar to the rate of NADH release ($k_5$). This, again, indicates that the elevated value of $k_{cat}$ may be caused by an increased degree of productive substrate binding. However, these kinetic parameters are directly comparable since a change in the regiopreference for B1 and B1F4 variants, as compared to the wild type.

Furthermore, wild type ADH-A displayed 78-fold preference in the oxidation of the monosubstituted vicinal diols for $(R)$-2 over $(S)$-2. Both B1 and B1F4, although displaying increases in $k_{cat}$ for $(S)$-2 (4 to 8-fold, respectively), retained the preference for the $(R)$-enantiomer (Table 15). The catalyzed reduction of 1,2-diketone (8) is notably efficient with all tested enzymes, a turnover of 160 s$^{-1}$ in the case of B1F4, a 2.5-fold increase as compared to wild type.

### Synthesis of 1-Phenylpropane-(1R,2S)-diol

The starting material were used in this work, $(1R,2S)$-3 was synthesized by employing epoxide hydrolase for potato, StEH1 (Figure 30A).

![Figure 30. A, epoxide hydroxylase (StEH1) catalyzed hydrolysis of (2S,3S)-epoxide resulted (1R,2S)-3 product. B, the oxidation of (1R,2S)-3 with ADH-A wild type predominantly reduced the benzylic “C-1” carbon of sec-alcohol to the corresponding (S)-6, and evolved variants (B1 and B1F4) predominantly reduced the “C-2” carbon of sec-alcohol to the corresponding (R)-7, and C, further coupled with aminotransferases, respectively.](image)

The wild type StEH1 is selectively converts the starting 1-phenylpropene-(1S,2S)-oxide into the corresponding 1-phenylpropane-(1R,2S)-diol.

### Regioselectivity Preference of ADH-A and Variants

In the oxidation reaction with $(1R,2S)$-3, wild type preferred C-1 and produced (S)-6 as major product with a ratio 3.6:1, both the (B1 and B1F4) variants insisted displayed regiopreference for C-2, and producing (R)-7 as major product with a ratio 3.4:1 and 4.5:1, respectively.
Figure 31. Characteristic peaks for methyl protons of (1R,2S)-3 (far-right) and two potential acyloin products; (S)-6 (center-right), (R)-7 (center-left), and 1-phenylpropane-1,2-dione, 8 (far-left). Beneath the peaks are 1H integrals after the oxidation reaction of (1R,2S)-3 by wild type or generated variants (B1 and B1F4) respectively. *, oxidation reaction this peak was not detected, and for simplification this peak was extracted for corresponding reduction reaction of 8. Similarly, reduction of 8, the potential acyloin products, 6 (center-right), 7 (center-left) were observed with an exception of 3, was not detected.

Table 16. ADH-A wild type and evolved variants (B1 and B1F4) oxidation reaction of (1R,2S)-3 and reduction reaction of corresponding cognate di-ketone 8, relative of formed products amounts was characterized by using 1H NMR. *, calculated from relative abundance of the distinguished methyl protons, as shown in Figure 27. a, not detected from oxidation reaction of (1R,2S)-3.

<table>
<thead>
<tr>
<th></th>
<th>Oxidation of (1R,2S)-3</th>
<th>Reduction of 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>B1</td>
</tr>
<tr>
<td>(1R,2S)-3</td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>(S)-6</td>
<td>72</td>
<td>70</td>
</tr>
<tr>
<td>(R)-7</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>a</td>
<td>-</td>
</tr>
</tbody>
</table>

However, in the reduction reaction with 1,2-diketone 8, wild type and both the variant preferred C-1 and produced acyloin 6 as the major product, respectively. The resulted respective products were characterized by using by 1H and 13C-NMR studies.
Molecular dynamics simulations
In an attempt to rationalize the experimental results of ADH-A we employed molecular dynamics (MD) simulations of selected enzyme-substrate complexes. The MD studies were conducted on the (R)-2, (S)-2, (1R,2S)-3 and 8 substrates in complex with wild type or variant enzymes.

The modeling of the wild type enzyme produced the following results: The 78-fold higher $k_{cat}$ with (R)-2, as compared to (S)-2, could be explained by the finding that the C-1 (benzylic carbon) hydride, in the case of (S)-2, points away from the nicotinamide C-4. Similarly, the case of (R)-1-phenylethanol, discussed in Paper I, where modeling by docking was performed. With (R)-2, the same C-1 hydrogen is directed towards the C-4 of NAD$^+$ at a distance of $\sim$3.1 Å, which can consider to be appropriate for catalysis. Modeling of (1R,2S)-3 resulted in clusters of conformations representing two different binding modes. Firstly, the substrate is oriented steered by the interactions with catalytic Zn$^{2+}$ in a conformation that could favor hydride transfer from C-1, with a distance approximately 2.7 Å to C-4 of the nicotinamide ring. In a second highly populated cluster, the distance in which the starting point involved coordination of the C-2 hydroxyl to the catalytic Zn$^{2+}$ ion, was similar to the previous case, but displaying relatively higher RMSD values for the substrate. In the case of the flat and rigid 1,2-diketone 8, no stable structure clusters were achieved within the time frame of the simulation that could describe catalytically competent enzyme-ketone complexes. This indicates that this substrate may bind in different binding modes inside the active site and the observed regioselectivity of the reduction of 8 is probably dependent subtle factors that influences the degree of productive binding modes of this substrate.

Conclusion:
The main objective of this work was to test the regioselectivity in the presence of a substrate with two secondary alcohol groups, and in parallel search for ADH-A variants that exhibited improved catalytic activity with aryl-substituted vicinal diols, such as 1-phenylpropane-(1R,2S)-diol ((1R,2S)-3) to enable efficient synthesis of corresponding acyloins. We could conclude that wild type ADH-A prefers oxidation of the benzylic carbon C-1 of (1R,2S)-3 generating (S)-6 as the major product, whereas the B1 (Y54G and L119Y) and B1F4 (F43T, Y54G, L119Y and F282W) variants displayed shifted regio-preference for oxidation of C-2 producing (R)-7 as major product. In the reduction of 1,2-diketone 8, both variants and the wild type preferred C-1 yielding acyloin 6 as major product. The shift in regioselectivity of the B1 and B1F4 variants can be attributed to the altered position of a tyrosine side chain caused by the double substitutions Y54G and L119Y, in the active site. The two additional mutations in B1F4 provides an additional increase in $k_{cat}$ for
(1R,2S)-3. As noted in previous work, parallel increases in the values of $k_{\text{cat}}$ and $K_M$ indicate a decrease of nonproductive binding of the substrate in the ternary enzyme-substrate complex.
Paper IV

In the oxidation reaction wild type ADH-A exhibits a 270-fold preference towards the (S)-enantiomer of 1-phenylethanol over the corresponding (R)-enantiomer. After applying laboratory evolution, we were able to switch the original stereoselectivity into a 7.4-fold enantiopreference for (R)-1-phenylethanol (Paper I). Similarly, the catalytic turnover with the vicinal 1,2-diols (R)-1-phenyl-1,2-ethanediol and 1-phenylpropane-(1R,2S)-diol were also improved by mining of ADH-A variant libraries (Papers II and III).

Considering the above described engineered ADH-A variants, selected for slightly diverse new traits (Figures 21 and Table 17), we wanted to investigate for possible promiscuity in substrate selectivity towards a spectrum of previously untested potential substrate alcohols and ketones (Figure 32 and 33).

Figure 32. ADH-A catalyzed chemical transformations; A, oxidation of rac-alcohols and B, reduction of ketones.

Table 17. Study involved wild type ADH-A and a selection of evolved enzyme variants (amino acid positions in the active site). #, from neighboring subunit of the dimer. °, indicates same as wild type ADH-A.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>39</th>
<th>43</th>
<th>54</th>
<th>119</th>
<th>282°</th>
<th>294</th>
<th>295</th>
<th>pdb ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>H</td>
<td>F</td>
<td>Y</td>
<td>L</td>
<td>F</td>
<td>Y</td>
<td>W</td>
<td>3vj7</td>
</tr>
<tr>
<td>A1</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>-</td>
</tr>
<tr>
<td>A2</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>5o8q</td>
</tr>
<tr>
<td>A2C3</td>
<td>Y</td>
<td>S</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>5o9f</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>Y</td>
<td>H</td>
<td>F</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>5o9d</td>
</tr>
<tr>
<td>B1</td>
<td>°</td>
<td>°</td>
<td>G</td>
<td>Y</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>5o3d</td>
</tr>
<tr>
<td>B1F4</td>
<td>°</td>
<td>T</td>
<td>G</td>
<td>Y</td>
<td>W</td>
<td>°</td>
<td>°</td>
<td>6fg0</td>
</tr>
<tr>
<td>C1</td>
<td>°</td>
<td>H</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>6ffx</td>
</tr>
<tr>
<td>C1B1</td>
<td>°</td>
<td>H</td>
<td>L</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>°</td>
<td>6ffx</td>
</tr>
</tbody>
</table>

From the results in Paper I, it was evident that W295 plays a decisive role in the enantiopreference for 1-phenylethanol. One question was therefore, if also other 1-methyl substituted alcohols could be affected similarly by the presence or absence of this residue. A more general issue was if these enzymes that had been originally selected for catalytic activities towards other alcohol substrates would display different overall substrate scopes as compared to the
wild type enzyme. Although selectivity in substrate recognition is considered an important component of enzyme catalysis, many enzymes are capable of catalyzing multiple reactions that transform also non-physiological substrates into products. Such lack of selectivity is referred to as enzyme promiscuity.


Kinetic Effect on Variants

We initially tested racemic alcohol mixtures (1–7), e.g. 1-phenylethanol (1), together with the corresponding ketones (8–14) for catalytic turnover. Following this first screening for activity scopes, the catalytic activities with pure enantiomers of alcohols 1, 3, 4 and 5 were determined for a more detailed understanding of the stereoselectivity in the oxidation reaction. We observed that many of the tested enzymes-substrate combinations displayed different degrees and types of cooperativity. The same substrate could elicit either positive, negative, or no cooperativity depending on the reacting enzyme.

Table 18. Steady-state kinetic parameters of ADH-A wild type and evolved variants. n°, Hill-coefficient, if n = 1, then the K0.5 becomes KM which is equal to the [S] that results in half maximum reaction velocity, and also kcat/K0.5.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>kcat (s⁻¹)</th>
<th>K0.5 (mM)</th>
<th>kcat/K0.5 (s⁻¹mM)</th>
<th>n°</th>
<th>kcat (s⁻¹)</th>
<th>K0.5 (mM)</th>
<th>kcat/K0.5 (s⁻¹mM)</th>
<th>n°</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>140±2</td>
<td>0.034±0.001</td>
<td>4300±6</td>
<td>1</td>
<td>140±1</td>
<td>0.061±0.002</td>
<td>2.300±0.002</td>
<td>1.4±0.05</td>
</tr>
<tr>
<td>A1</td>
<td>90±2</td>
<td>0.55±0.06</td>
<td>170±10</td>
<td>1</td>
<td>82±2</td>
<td>1.6±0.2</td>
<td>50±5</td>
<td>1.4±0.2</td>
</tr>
<tr>
<td>A2</td>
<td>8×2</td>
<td>0.42±0.03</td>
<td>200±10</td>
<td>1</td>
<td>49±1</td>
<td>1.7±0.1</td>
<td>28±3</td>
<td>1.3±0.1</td>
</tr>
<tr>
<td>A2C3</td>
<td>89±4</td>
<td>1.1±0.2</td>
<td>78±9</td>
<td>1</td>
<td>61±2</td>
<td>1.8±0.2</td>
<td>33±4</td>
<td>1.4±0.2</td>
</tr>
<tr>
<td>A2C2B1</td>
<td>50±1</td>
<td>0.73±0.1</td>
<td>79±10</td>
<td>1</td>
<td>73±3</td>
<td>3.0±0.6</td>
<td>2.4±0.4</td>
<td>1</td>
</tr>
<tr>
<td>B1</td>
<td>85±1</td>
<td>0.21±0.008</td>
<td>390±10</td>
<td>1.3±0.05</td>
<td>100±4</td>
<td>0.61±0.04</td>
<td>160±9</td>
<td>1</td>
</tr>
<tr>
<td>B1F4</td>
<td>76±1</td>
<td>0.19±0.02</td>
<td>390±10</td>
<td>1</td>
<td>91±3</td>
<td>1.1±0.2</td>
<td>84±10</td>
<td>1</td>
</tr>
<tr>
<td>C1</td>
<td>140±2</td>
<td>0.13±0.006</td>
<td>1100±50</td>
<td>1.4±0.07</td>
<td>220±6</td>
<td>0.25±0.02</td>
<td>870±60</td>
<td>1</td>
</tr>
<tr>
<td>C1B1</td>
<td>170±3</td>
<td>0.14±0.01</td>
<td>1200±70</td>
<td>1</td>
<td>200±3</td>
<td>0.44±0.02</td>
<td>460±20</td>
<td>1</td>
</tr>
</tbody>
</table>
For example, the wild type displayed strong positive cooperativity with a Hill-coefficient of 1.8 in the oxidation of 4-phenyl-2-butanol (4). Similar behavior was observed with B1 (n=1.3), and C1 (n=1.4), whereas A1, A2, A2C3 and A2C2B1 displayed non-cooperative Michaelis-Menten saturation kinetics (n=1). The wild type reduction of corresponding cognate ketone, 4-phenyl-2-butane (11) also displayed strong positive cooperativity (n=1.8). In this case also A1 (n=1.4), A2 (n=1.3) and A2C3 (n=1.4), showed positive cooperativity, whereas A2C2B1, B1, B1F4, C1 and C1B1 displayed 'normal' Michaelis-Menten saturation kinetics. Thus, A1, A2 and A2C3 switched from non-cooperative behavior to positive cooperativity in substrate dependency, when moving from alcohol oxidation to ketone reduction. The reverse was observed for variants B1 and C1 which switched from positive cooperativity in alcohol oxidation to non-cooperative saturation kinetics, in the ketone reduction (Table 18).

Furthermore, we tested the respective enantiomers of alcohol 1, 3, 4 and 5, to validate the enantiopreference. As described in the Paper I, the replacement of W295A substitution in variants A1, A2, A2C1 and A2C2B1 allows the rotation of substrates in the active site, thus facilitating the increase in productive binding with (R)-1 but causing a parallel decrease in activity with (S)-1, presumably due to new nonproductive binding modes in the enlarged active site. Apparently, this behavior extends to alcohols 4 and 5 with increases in kcat, with the (R)-enantiomer and decreased activities with the S-enantiomers. This shift towards (R)-preference is, however, not observed in the products following reduction of the corresponding ketones, tested with variants A2C2 and A2C2B1. The alcohol products are essentially racemic mixtures. This is in agreement with the earlier results in Paper I where the W295A containing variants favored reduction of (R)-1-phenylethanol, but in reduction of acetophenone produced racemates of both enantiomers. In the wild type catalyzed reduction of acetophenone, or ketones 11 and 12 studied here, the (S)-alcohol are exclusively produced, a direct reflection of the selectivity of the corresponding alcohol oxidations. (Table 19).

<table>
<thead>
<tr>
<th>Table 19. Enantioselectivity in oxidation and reduction of alcohol and ketone. a, ratios (kcat/Km)major/(kcat/Km)minor. b, data is extracted after chiral HPLC separation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product enantiomeric ratio, (R) : (S)</td>
</tr>
<tr>
<td><strong>Substrate</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>wild type</td>
</tr>
<tr>
<td>A2C3</td>
</tr>
<tr>
<td>A2C2B1</td>
</tr>
</tbody>
</table>
Conclusion:
The main objective of this work\textsuperscript{124} was to study the diverse selectivity traits in engineered ADH-A variants from A–C library, we exploited possible promiscuities in substrate selectivity towards a spectrum of previously untested potential alcohol and ketone substrates. It is clearly evident that the W295A substitution plays a decisive role in enantioselective outcome for oxidation of alcohols 4 and 5, presumably caused by decreases in nonproductive binding modes resulting in elevated values of $k_{cat}$. The wild type enzyme displays strong preference for the (S)-enantiomers in the oxidation of alcohols 1, 3, 4 and 5, and produces the same product enantiomers after reduction of the cognate ketones.
Concluding Remarks and Future Perspective

The main objective of this work has been the engineering of alcohol dehydrogenase-A (ADH-A) from *Rhodococcus ruber* DSM 44541 and the subsequent characterization of resulting enzyme variants, to assess their potential as biocatalysts of chemical transformations. ADH-A tolerates water miscible organic solvents, accepts a relatively wide range of arylated sec-alcohols and vicinal 1,2-diols and the corresponding ketones as substrates in redox reactions.

The engineering efforts have altered active site cavity by employing saturation mutagenesis. The evolved variants were selected for new substrate scopes, including altered stereo- and regioselectivities. Since arylated ketones, α-hydroxy ketones and their corresponding alcohols are important building blocks in the synthesis of fine chemicals, pharmaceuticals and natural products, one objective of this work has been to generate ADH-A variants that could act as putative biocatalysts for the production of aryl-substituted ketones and asymmetric alcohols.

The oxidation of vicinal 1,2-diols involves the removal of an asymmetric center into a prochiral ketone. However, the corresponding acyloin is an attractive building block for the production of important carbon–carbon bonded derivatives, such as chiral auxiliaries, natural products, fine chemicals and pharmaceuticals. Therefore, in this study we tried to improve the catalytic activity of ADH-A with vicinal 1,2-diols, such as (R)-1-phenyl-1,2-ethanediol, by generating variants by the same approach as described in Paper I. ADH-A variants of improved catalytic activities may be directly employed in multi-step enzyme catalyzed reactions, thus, a starting reactant can be cheaper racemic styrene oxide, which can be transformed into a more valuable acyloin.

In the oxidation reaction wild type ADH-A exhibits a 270-fold preference towards the (S)-enantiomer of 1-phenylethanol over the corresponding (R)-enantiomer. After applying laboratory evolution, we were able to switch the original stereoselectivity into a 7.4-fold enantiopreference for (R)-1-phenylethanol (Paper I). Similarly, the catalytic turnover with the vicinal 1,2-diols (R)-1-phenyl-1,2-ethanediol and 1-phenylpropane-(1R,2S)-diol were also improved by mining of ADH-A variant libraries (Papers II and III).
as compared to (S)-1-phenylethanol. As described in Paper II, ADH-A catalyzed oxidation of vicinal 1,2-diols results in the removal of an asymmetric center into a prochiral ketone. Therefore, in this study we focused on generating variants that displayed improved regioselectivity in oxidation with a (1R,2S)-1-phenyl-1,2-propanediol, by the same approach as described in Paper I. ADH-A variants of improved catalytic activities may be directly employed in multi-step enzyme catalyzed reactions, thus, a starting reactant can be cheaper racemic styrene oxide, which can be transformed into a more valuable acyloins, e.g. to produce chiral amino alcohols.

Figure 34. A, epoxide hydroxylase (StEH1) catalyzed hydrolysis of (2S,3S)-epoxide resulted 1-phenylpropane-(1R,2S)-diol. B, the oxidation of (1R,2S)-product with predominantly reduced the benzylic “C-1” or “C-2” carbon of sec-alcohol to the corresponding (S)- or (R)-acyloin by evolved variants, and C, further coupled with aminotransferases, respectively.

Furthermore, possible substrate promiscuity in engineered enzyme variants has also been addressed.
Svensk Sammanfattning

Enzymer är biomolekyler byggda av aminosyror och katalyserar de kemiska transformationerna i en cell. Enzymer är naturligt stereoselektiva, biologiskt nedbrytbara, miljövänliga och kan katalysera reaktioner i vattenbaserade lösningar och i rumstemperatur. Tack vare dessa fördelar, så har intresset ökat för att använda enzymer som biokatalysatorer vid kemiska transformationer istället för traditionella kemiska syntesstrategier. I de fall naturliga enzymer inte kan katalysera de önskade kemiska transformationerna, så kan de modifieras genom riktad evolution eller andra proteinmodifierande tekniker.

Då enzymer är genetiskt kodade så kan de optimeras för önskade färdigheter, som till exempel substratelektivitet eller ökad katalytisk förmåga. Med dessa fördelar i åtanke, och samtidigt ta hänsyn till de syntetiska och industriella användningsområdena, så har vi utnyttjat alkohol dehydrogenas-A (ADH-A) från *Rhodococcus ruber* DSM 44541 som ett studieexempel för att ta fram nya katalytiska egenskaper. ADH-A tål vattenlösliga organiska lösningsmedel, accepterar ett antal aromatiska sekundära alkoholer och ketoner som substrat och är därför en potentiellt användbar biokatalysator för asymmetrisk syntes av organiska föreningar.

![Diagram av katalys av (R) selektiv kinetiskupplösning](image)

*Figure 35. Exempel på katalys av (R) selektiv kinetiskupplösning.*

Det presenterade forskningsarbetet i denna avhandling har huvudsakligen fokuserat på modifieringen av ADH-A och karaktäriseringen av de enzymvarianten som framkommit. Målet med enzymmodifieringarna har varit att ändra enzymets substratspektra, inklusive dess stereo- och regioselektivitet.
Dessutom har möjlig substratpromiskuitet i de modifierade varianterna studerats. Sammanfattningsvis,


Figure 36. Exempel på katalys.


Figure 37. Exempel regioselektivitet.

iv). Artikel IV fokuserade på att studera de möjliga effekterna på substratpromiskuitet hos modifierade varianter jämfört mot vildtypszysetet när de testades med ett spektrum av potentiella, tidigare ej testade, substrat.
Acknowledgements

This thesis would not have possible without the support and help from my colleagues, collaborators, friends and family. Therefore, I would like to express my sincere gratitude to everyone.

****First and foremost, my Guru & advisor: thank you Prof. Mikael Widersten for having faith in me and providing for me this opportunity to do PhD in your team in Uppsala. I sincerely appreciate you Micke, for giving me the interesting FucO, ADH-A, FSA and synthetic organic chemistry projects for research explorations (oops, just missed the old work horse, the potato StEH1®. Thank you for allowing me to work independently & act independently. Furthermore, you encouraged me to do all sorts of diverse things “if I can say so – this will sound like Induced–Promiscuity”. Thanks a lot for the team meeting, fika, coffee/tea break and all sorts of discussions. Most importantly, guidance and sharing the knowledge to grow as a person with humility and empathy. ***Again, thank you so much Micke, for helping me to write this thesis, corrections and proof reading.

*Thank you Dr. Eszter Borbas for all the support.

*I would like to thank MW-Team members past and present: Emil Hamnevik, Åsa Janfalk Carlsson, Huan Ma, Gina Chukwu, Cecilia Blikstad and Derar Smadi, Giuseppe, Dirk Maurer, Anika and Khyati.

*I sincerely thank all my collaborators and their invaluable contribution for ADH-A, FSA, FucO and synthetic projects: Dr. Doreen Dobritzsch, Prof. Lyn Kamerlin, Prof. Rikkert K. Wierenga, Prof. Thomas Norberg and Prof. Jan Kihlberg.

*Especially, I would like to thank all the senior biochemistry colleagues: Prof. Helena Danielson, Prof. Gunnar Johansson, Dr. Ylva Ivarsson and Dr. Erik Marklund.

*I would like to thank all fellow colleagues in biochemistry section: Emil, Eldar, Gina, Anika, Khyati, Sandra, Daniela, Edward, Helena, Mostafa, Helena, Vladimir, Giulia, Gun, Dirk, Leandro, Ali, Caroline, Susanna, Johanna, Dilip, Erika, Joana Filipa, Maxim, Yashraj, Anil and Malin.
*Thank you, I am so grateful for your contribution; ADH-A work: Emil Hamnevik and Dirk Maurer. FSA work: Huan Ma and Derar Smadi. FucO work: Shruthi Shridhar.

*I would like to thank everyone at the Chemistry – BMC, who helped me: Adolf, Johan, Rikard, Ergun, Yoseph, Karthik, Mohit, Jagadish, Fredric, Matic, Fabio, Mate, Xiong, Scott, Hermina, Daniel, Stefan, Duc Duy, Luk, and Chi.

*Thank you, Eva, Hanna, Johanna and Gunnar for the support.

*Thank you, Emil Hamnevik, Eldar Abdurakmanov and Srinivas Akula, Yashraj Kulakarni.

*Thank you so much, Kerri Sandell and Jesper Andersson for your kindness and generosity.

*****Finally, thanks to my family members, mom, dad, brothers and sister-in-law, niece, cousins and my wife for their love, care and ∞ support*****

***Thank You All***
References


(2) Intergovernmental Panel on Climate Change; United Nations; WHO. *Special Report on Global Warming of 1.5 °C A Summary for Policymakers*, 2019, 1–24.


(77) Hill, A. V. A New Mathematical Treatment of Changes of Ionic Concentration in Muscle and Nerve under the Action of Electric Currents,


(109) Olson, L. P.; Luo, J.; Almarsson, O.; Bruce, T. C. Mechanism of Aldehyde Oxidation Catalyzed by Horse Liver Alcohol Dehydrogenase. *Biochemistry* 1996, 35, 9782–9791.

(110) Eklund, H.; Nordström, B.; Zepezauer, E.; Söderlund, G.; Ohlsson, I.; Bouwe, T.; Söderberg, B.-O.; Tapia, O.; Brändén, C.-I.; Åkeson, Å.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1872

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”).

Distribution: publications.uu.se
urn:nbn:se:uu:diva-395527